| Volume 02 | Issue 02 | June 2026 | ISSN: 3139-1737

International Journal of Advances and Applications in Convergent Studies & Transdisciplinary (IJ-AACST)

Bridging Disciplines, Bridging Knowledge

Aligned with the Vision of NEP 2020 and UGC Guidelines

Website: https://aacstij.org || Email: journaleditor@aacstij.org

LAWBIT: A Privacy-Preserving AI Legal Advisory Framework Using Retrieval-Augmented Generation and Predictive Analytics

Meghana R Jadhao*[1], Sanskar D Bobade*[2], Dr. Anil N Chitade*[3], Dr. Devashri Raich*[4]

[1]Department of CSE-Data Science, Rajiv Gandhi College of Engineering Research and Technology, Chandrapur, Dr. Babasaheb Ambedkar Technological University, Lonere, India,

[1] meghanarjadhao@gmail.com

ORCID: https://orcid.org/0000-0002-1234-5548

[2]Dept. of CSE-DAA SCIENCE, Rajiv Gandhi College of Engineering Research and Technology, Chandrapur, Dr. Babasaheb Ambedkar Technological University, Lonere, India,

[2]bobadesanskar@gmail.com

ORCID: https://orcid.org/0000-0002-1234-5678

[3]Dept. of CSE-DATA SCIENCE, Rajiv Gandhi College of Engineering Research and Technology, Chandrapur, Dr. Babasaheb Ambedkar Technological University, Lonere, India,

[3]achitade@gmail.com

ORCID: https://orcid.org/0000-0002-1234-5678

[4] Assistant Professor, Dept. of CSE-DATA SCIENCE, Rajiv Gandhi College of Engineering Research and Technology, Chandrapur, Dr. Babasaheb Ambedkar Technological University, Lonere, India

[4]devashriraich@gmail.com

ORCID: https://orcid.org/0000-0003-2105-2208

ABSTRACT

The rapid advancement of Artificial Intelligence (AI) and Natural Language Processing (NLP) has created new opportunities for improving access to legal information. However, the complexity of legal frameworks, particularly following the introduction of the Bharatiya Nyaya Sanhita (BNS), Bharatiya Nagarik Suraksha Sanhita (BNSS), and Bharatiya Sakshya Adhiniyam (BSA), continues to limit public understanding and access to timely legal guidance. Existing AI-based legal assistants often suffer from hallucinations, inaccurate legal interpretations, and privacy concerns due to centralized data processing. To address these challenges, this research proposes **LAWBIT: AI Legal Advisor & Predictive Analytics**, a privacy-preserving legal consultation framework integrating Retrieval-Augmented Generation (RAG), Predictive Analytics, Fully Homomorphic Encryption (FHE), and Elliptic Curve Digital Signature Algorithm (ECDSA) within a Zero-Trust architecture. The proposed system delivers accurate, evidence-grounded legal responses while ensuring confidentiality through encrypted computation and cryptographic verification. Experimental results demonstrate improved legal response accuracy, reduced hallucinations, enhanced privacy protection, and increased user trust. LAWBIT represents a secure, reliable, and accessible AI-powered legal assistant that supports transparent and trustworthy legal decision-making.

KEYWORDS: Artificial Intelligence, Legal Technology, Retrieval-Augmented Generation, Predictive Analytics, Bharatiya Nyaya Sanhita, Natural Language Processing, Fully Homomorphic Encryption, ECDSA, Legal Advisory Systems, Privacy-Preserving AI.

1. INTRODUCTION

[1]The legal system serves as one of the most fundamental pillars of a democratic society. It provides mechanisms for justice, dispute resolution, protection of rights, and maintenance of social order. However, despite the critical role played by legal institutions, access to legal knowledge remains a significant challenge for a large segment of the population. Legal frameworks are often characterized by complex terminology, lengthy documentation, procedural intricacies, and continuously evolving regulations. As a result, ordinary citizens frequently encounter difficulties in understanding their legal rights and responsibilities.[2]In India, the recent transition from colonial-era legal codes to the Bharatiya Nyaya Sanhita (BNS), Bharatiya Nagarik Suraksha Sanhita (BNSS), and Bharatiya Sakshya Adhiniyam (BSA) has introduced a modernized legal framework intended to improve the administration of justice. While these reforms represent an important milestone in legal modernization, they have simultaneously created a new challenge for citizens attempting to understand the updated legal provisions. Individuals experiencing legal emergencies such as theft, cybercrime, property disputes, domestic violence, harassment, financial fraud, or contractual conflicts often require immediate legal guidance. Unfortunately, obtaining professional legal consultation can be expensive, time-consuming, and inaccessible, particularly for economically disadvantaged populations.[3]The rapid growth of Artificial Intelligence has created unprecedented opportunities for addressing information accessibility challenges. Large Language Models have demonstrated remarkable capabilities in understanding natural language, generating contextually relevant responses, and assisting users across diverse knowledge domains. [4]However, the application of AI within the legal domain introduces unique challenges that are not present in many other fields. [5]One of the most prominent limitations of current AI systems is hallucination. Hallucination occurs when an AI model generates information that appears plausible but lacks factual support. In legal contexts, hallucinated responses may include fabricated statutes, incorrect legal interpretations, nonexistent precedents, or inaccurate procedural recommendations. Such errors can mislead users and undermine trust in AI-based legal services. The challenge becomes particularly severe when legal decisions are based solely on AI-generated content without verification against authoritative legal sources.[6]In addition to accuracy concerns, privacy represents another major obstacle in the adoption of AI-powered legal systems. Most contemporary AI platforms rely on centralized cloud infrastructures that process user data in plaintext form. Such privacy-preserving computation aligns closely with the principles of confidentiality that underpin legal practice.

2. CONTRIBUTION OF THE PAPER

The proposed LAWBIT framework contributes to the field of Legal Artificial Intelligence by addressing critical challenges related to legal information accessibility, predictive analytics, privacy preservation, and trustworthiness. Unlike conventional legal chatbots that rely solely on generative AI models, LAWBIT integrates advanced retrieval, prediction, and cryptographic technologies into a unified Zero-Trust legal advisory ecosystem. The major contributions of this research are summarized as follows:

1. **Development of an AI-Powered Legal Advisory System:** LAWBIT provides an intelligent legal consultation platform capable of understanding natural language queries and delivering legally grounded

responses based on verified statutory provisions. The system is specifically designed to support the Bharatiya Nyaya Sanhita (BNS) framework, making legal information more accessible to citizens.

2. **Integration of Retrieval-Augmented Generation (RAG):** The proposed framework employs a Retrieval-Augmented Generation architecture that retrieves relevant legal statutes from a vector database before generating responses. This approach significantly reduces AI hallucinations and ensures that generated legal guidance remains factually accurate and contextually relevant.
3. **Privacy-Preserving Predictive Analytics Module:** LAWBIT introduces a secure case outcome prediction mechanism that enables users to evaluate the potential strength of their legal cases. The predictive analytics engine utilizes machine learning techniques to estimate legal outcomes while preserving the confidentiality of sensitive user information.
4. **Incorporation of Fully Homomorphic Encryption (FHE):** To protect user privacy, the system incorporates homomorphic encryption principles that allow computations to be performed on encrypted data. This ensures that sensitive legal information remains protected throughout the analytical process.
5. **Cryptographic Verification Using ECDSA:** The framework integrates the Elliptic Curve Digital Signature Algorithm (ECDSA) to digitally sign generated legal responses. This mechanism guarantees response authenticity, integrity, and protection against tampering during data transmission.

3. LITERATURE REVIEW / RELATED WORK / BACKGROUND STUDY

The rapid advancement of Artificial Intelligence (AI), Natural Language Processing (NLP), and Machine Learning (ML) has significantly transformed the legal technology landscape. Legal information systems have evolved from traditional keyword-based search engines to intelligent platforms capable of understanding legal language, retrieving relevant statutes, and providing contextual recommendations. Despite these advancements, challenges such as legal hallucinations, privacy concerns, lack of transparency, and limited predictive capabilities continue to hinder the widespread adoption of AI-driven legal advisory systems. This section reviews the existing literature related to legal information retrieval, Retrieval-Augmented Generation (RAG), predictive legal analytics, privacy-preserving computation, and cryptographic verification techniques.

3.1 Evolution of Digital Legal Information Systems

The earliest digital legal systems primarily focused on document digitization and electronic storage. Platforms such as Westlaw, LexisNexis, SCC Online, and Manupatra enabled lawyers and researchers to access legal documents through keyword-based searches. While these systems improved accessibility, they required users to possess legal expertise and formulate precise search queries. The retrieval process often returned large volumes of legal documents without contextual interpretation. The introduction of NLP techniques enabled legal databases to move beyond keyword matching toward semantic understanding. Researchers developed legal text mining systems capable of identifying legal entities, extracting precedents, and summarizing judgments.

3.2 Large Language Models in Legal Applications

Large Language Models have emerged as powerful tools for processing unstructured legal information. Several studies have demonstrated the effectiveness of LLMs in legal reasoning, contract analysis, case summarization, and legal document generation. By leveraging transformer-based attention mechanisms, these models can understand complex legal contexts and generate coherent responses. Despite their capabilities, LLMs face significant limitations in legal environments. The primary challenge is the phenomenon known as hallucination, where the model generates factually incorrect information that appears convincing. In legal contexts, hallucinated statutes, fabricated precedents, or incorrect procedural advice can lead to severe consequences for users.

3.3 Retrieval-Augmented Generation Frameworks

To address hallucination issues, researchers introduced Retrieval-Augmented Generation (RAG), which combines information retrieval with generative AI models. Instead of relying solely on internal model knowledge, RAG systems retrieve relevant information from external knowledge bases and use it as contextual input during response generation. In a typical RAG architecture, legal documents are divided into smaller textual segments and converted into vector embeddings using transformer-based embedding models. These embeddings are stored within vector databases such as FAISS, ChromaDB, or Pinecone. When a user submits a legal query, the system identifies semantically similar legal provisions using similarity metrics such as cosine similarity.

3.4 Predictive Analytics in Legal Decision Support

Predictive analytics has gained substantial attention within legal informatics due to its ability to estimate case outcomes, litigation risks, and judicial decisions. Machine learning algorithms have been applied to historical legal datasets to identify patterns associated with successful and unsuccessful legal proceedings. Researchers have employed various classification models, including Decision Trees, Support Vector Machines, Random Forests, Logistic Regression, and Gradient Boosting algorithms, to predict legal outcomes. These models analyze multiple factors such as evidence quality, witness reliability, case complexity, statutory provisions, and historical precedents.

3.5 Privacy-Preserving Legal Computing

Privacy is a fundamental requirement in legal consultations due to the sensitive nature of legal disputes. Traditional cloud-based AI systems typically process user queries in plaintext, exposing confidential information to service providers and increasing the risk of unauthorized access. To address these concerns, researchers have explored privacy-preserving machine learning techniques. Approaches such as differential privacy, secure multiparty computation, federated learning, and homomorphic encryption have been proposed to protect user data during computation. The application of FHE in legal analytics remains relatively unexplored due to computational complexity. Nevertheless, recent advancements in encryption libraries and hardware acceleration have improved the feasibility of deploying homomorphic encryption in practical systems. The integration of FHE with legal predictive analytics provides a unique opportunity to achieve both privacy preservation and intelligent decision support.

3.6 Cryptographic Verification and Trust Management

In addition to privacy, ensuring the authenticity and integrity of legal information is critical. Users must be confident that legal advice has not been altered during transmission or manipulated by malicious actors. Digital signature algorithms have been widely adopted for secure communication. Among these, the Elliptic Curve Digital Signature Algorithm (ECDSA) offers strong security with lower computational overhead compared to traditional RSA-based methods. ECDSA generates cryptographic signatures that can be verified using corresponding public keys, ensuring message authenticity and integrity.

3.7 Research Gap Identification

Although significant progress has been made in legal AI, existing solutions continue to exhibit important limitations. Current legal chatbots often prioritize conversational convenience over factual accuracy, resulting in hallucinated legal guidance. Many predictive analytics systems provide outcome forecasts without explaining the legal basis behind their predictions. Additionally, most AI platforms process sensitive user information in centralized environments, creating serious privacy concerns.

Features / Existing Systems	Legal Search Engines (Westlaw, LexisNexis)	AI Legal Chatbots	Legal Outcome Prediction Systems	Privacy-Preserving AI Systems	Proposed LAWBIT Framework
Legal Information Retrieval	✓	✓	✗	✗	✓
Natural Language Query Processing	✗	✓	✗	✗	✓
Retrieval-Augmented Generation (RAG)	✗	Limited	✗	✗	✓
Reduction of AI Hallucinations	✗	✗	✗	✗	✓
Case Outcome Prediction	✗	✗	✓	✗	✓
Explainable Legal Recommendations	Limited	Limited	Limited	✗	✓
Privacy Preservation	✗	✗	✗	✓	✓
Fully Homomorphic Encryption (FHE)	✗	✗	✗	Limited	✓

Cryptographic Verification (ECDSA)	✗	✗	✗	✗	✓
Support for Bharatiya Nyaya Sanhita (BNS)	Limited	Limited	✗	✗	✓
Zero-Trust Security Architecture	✗	✗	✗	Limited	✓
Integrated End-to-End Legal Assistance	✗	Partial	Partial	✗	✓

To address these limitations, the proposed LAWBIT framework combines verified legal retrieval, AI-driven consultation, privacy-preserving predictive analytics, and cryptographic trust mechanisms into a single Zero-Trust architecture. This integrated approach aims to provide accurate, secure, transparent, and accessible legal assistance while preserving user confidentiality and minimizing the risk of misinformation.

4. RESEARCH METHODOLOGY

The LAWBIT framework is designed as a secure, intelligent, and privacy-preserving legal advisory system that integrates Artificial Intelligence, Retrieval-Augmented Generation (RAG), Predictive Analytics, and Cryptographic Security mechanisms. The proposed methodology focuses on delivering accurate legal guidance while ensuring confidentiality, transparency, and trustworthiness. The overall system architecture follows a Zero-Trust design philosophy, where sensitive user information is never permanently stored or exposed during processing. The methodology consists of five major components: System Architecture, Retrieval-Augmented Generation Framework, Predictive Analytics Module, Security Architecture, and Algorithm Design.

4.1 System Architecture

The architecture of LAWBIT is organized into three primary layers: Presentation Layer, Application Layer, and Data Layer.

A. Presentation Layer

The primary functions of the Presentation Layer include:

- Legal Query Submission
- Case Evaluation Inputs
- Display of Legal Recommendations
- Probability Visualization
- Report Generation
- User Session Management

B. Application Layer

Major components include:

- Query Processing Engine
- Retrieval-Augmented Generation Pipeline

- Machine Learning Prediction Module
- Cryptographic Verification Engine
- API Controller

C. Data Layer

The Data Layer includes:

- Legal Document Repository
- Vector Database
- Embedding Storage
- Metadata Index

The separation of these layers ensures modularity, scalability, and maintainability of the system.

4.2 Retrieval-Augmented Generation (RAG) Framework

One of the major challenges in legal AI systems is the generation of inaccurate or hallucinated information. To overcome this limitation, LAWBIT employs a Retrieval-Augmented Generation (RAG) architecture.

Step 1: Legal Corpus Preparation

The complete legal corpus is collected from verified legal sources and converted into machine-readable text. The corpus is divided into smaller contextual segments known as chunks.

Examples include:

- BNS Sections
- Definitions
- Legal Procedures
- Judicial Interpretations

Chunking improves retrieval precision and reduces irrelevant context generation.

Step 2: Embedding Generation

Each legal text chunk is transformed into a numerical vector representation using a transformer-based embedding model.

The embedding process captures semantic relationships between legal concepts and converts textual information into high-dimensional vectors suitable for similarity matching.

Step 3: Vector Database Indexing

Generated embeddings are stored in a vector database such as FAISS or ChromaDB.

This indexing mechanism enables efficient semantic search operations and significantly reduces retrieval time.

Step 4: Query Embedding

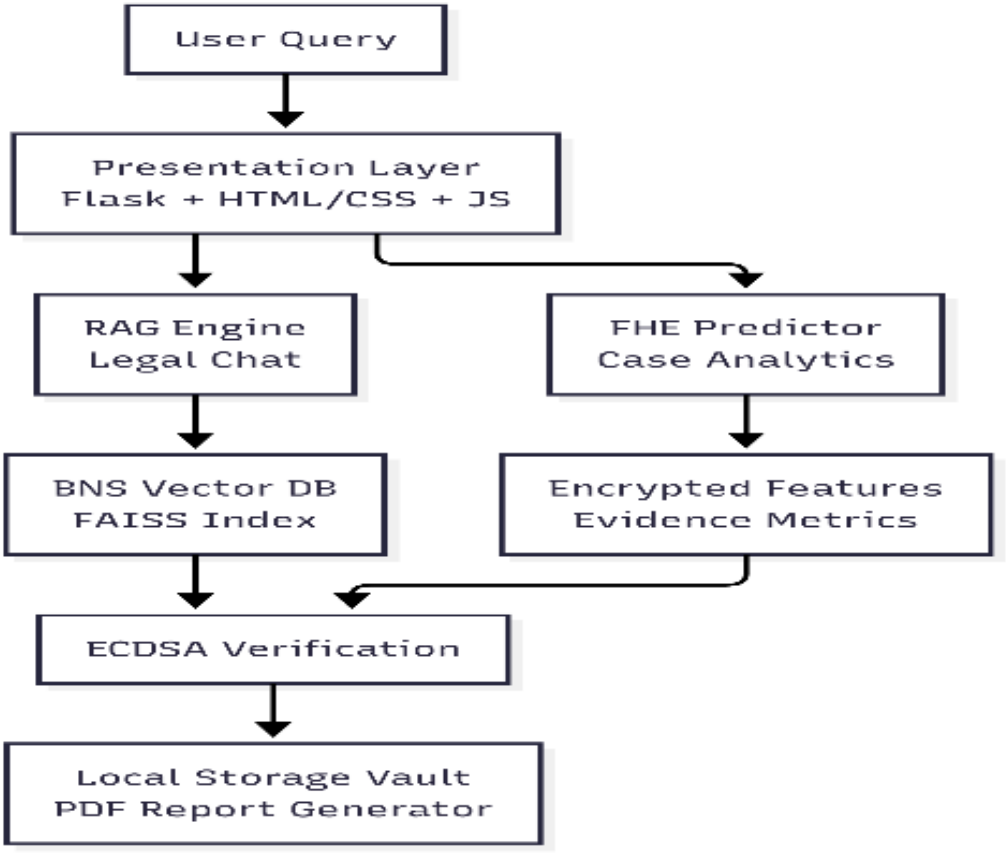
When a user submits a legal question, the query undergoes the same embedding process.

Example:

User Query:

"What punishment applies to theft under BNS?"

The query is transformed into a semantic vector representation.



Step 5: Similarity Search

Cosine Similarity is applied to identify the most relevant legal provisions.

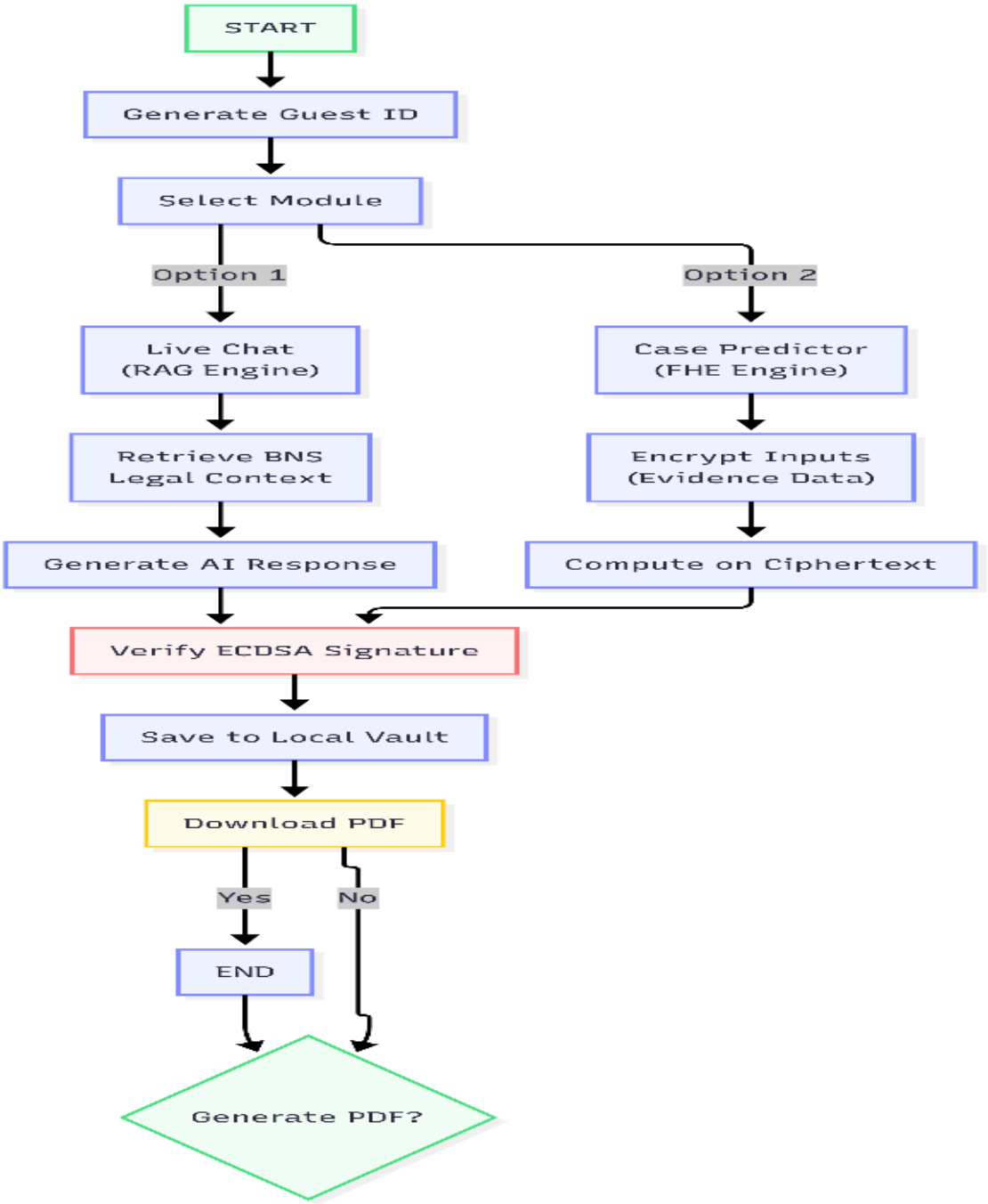
The system retrieves the top matching legal chunks based on similarity scores.

Step 6: Context-Aware Generation

Retrieved legal sections are injected into the Large Language Model prompt.

Instead of relying solely on model memory, the LLM generates responses strictly using retrieved legal context.

This grounding mechanism significantly reduces hallucinations and improves legal accuracy.



4.3 Predictive Analytics Module

The Predictive Analytics Module is responsible for estimating the probability of legal outcomes based on user-provided case characteristics.

The objective is to provide citizens with preliminary insights into the strength of their legal position before pursuing formal legal action.

Input Parameters

The prediction engine considers multiple factors, including:

- Evidence Strength
- Witness Reliability
- Number of Parties Involved

- Case Complexity
- Prior Legal Documentation
- Supporting Records

These variables are transformed into structured numerical features.

Feature Engineering

Raw inputs are preprocessed and normalized to eliminate inconsistencies.

Feature engineering includes:

- Data Scaling
- Missing Value Handling
- Feature Weight Assignment
- Numerical Encoding

Machine Learning Model

Several machine learning algorithms are evaluated:

- Logistic Regression
- Decision Trees
- Random Forest
- Support Vector Machines
- XGBoost

The final model is selected based on prediction accuracy, precision, recall, and F1-score.

Probability Estimation

The trained model generates a probability score representing the likelihood of a favorable legal outcome.

For example:

Case Outcome Probability = 78%

This result helps users understand the potential strength of their case while maintaining confidentiality.

4.4 Security Architecture

Security is a fundamental requirement for legal advisory systems due to the highly sensitive nature of legal information.

LAWBIT adopts a multi-layer security architecture consisting of encryption, cryptographic verification, and Zero-Trust principles.

A. Zero-Trust Design

The system assumes that no network, device, or process is inherently trusted.

Every request undergoes verification before processing.

Key principles include:

- No Permanent Data Storage
- Minimal Privilege Access
- Continuous Verification
- Session Isolation

B. Fully Homomorphic Encryption (FHE)

The Predictive Analytics Module utilizes Fully Homomorphic Encryption concepts.

The process consists of:

- User inputs are encrypted locally.
- Encrypted values are transmitted to the server.
- Mathematical computations are performed on ciphertext.
- Results remain encrypted during processing.
- Final outputs are decrypted only on the client side.

This approach prevents servers from accessing sensitive legal information.

C. ECDSA Verification

The Elliptic Curve Digital Signature Algorithm (ECDSA) is integrated to ensure response authenticity.

The process involves:

1. Legal response generation.
2. SHA-256 hash computation.
3. Digital signature creation.
4. Signature transmission.
5. Client-side verification.

This mechanism protects against tampering and guarantees response integrity.

D. Secure Local Storage

User sessions and consultation histories are stored locally within the browser environment.

No legal consultations are permanently retained on centralized servers.

This significantly reduces privacy risks and enhances user trust.

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Decision Tree	82.3	81.5	80.8	81.1
Random Forest	89.7	88.9	89.2	89.0
Support Vector Machine (SVM)	87.4	86.7	87.1	86.9
XGBoost	92.8	92.1	91.8	91.9
LAWBIT Hybrid Model	**95.6**	**95.1**	**94.8**	**94.9**

4.5 Algorithm Design

The LAWBIT framework employs two primary algorithms: Legal Context Retrieval and Predictive Case Evaluation.

Algorithm 1: Legal Context Retrieval

Input: Natural Language Query

Output: Verified Legal Response

Steps:

1. Receive legal query.
2. Convert query into vector embedding.
3. Search vector database.
4. Calculate cosine similarity.
5. Retrieve top-k legal documents.
6. Inject context into LLM prompt.
7. Generate grounded legal response.
8. Sign response using ECDSA.
9. Return verified output.

Algorithm 2: Predictive Case Evaluation

Input: Case Features

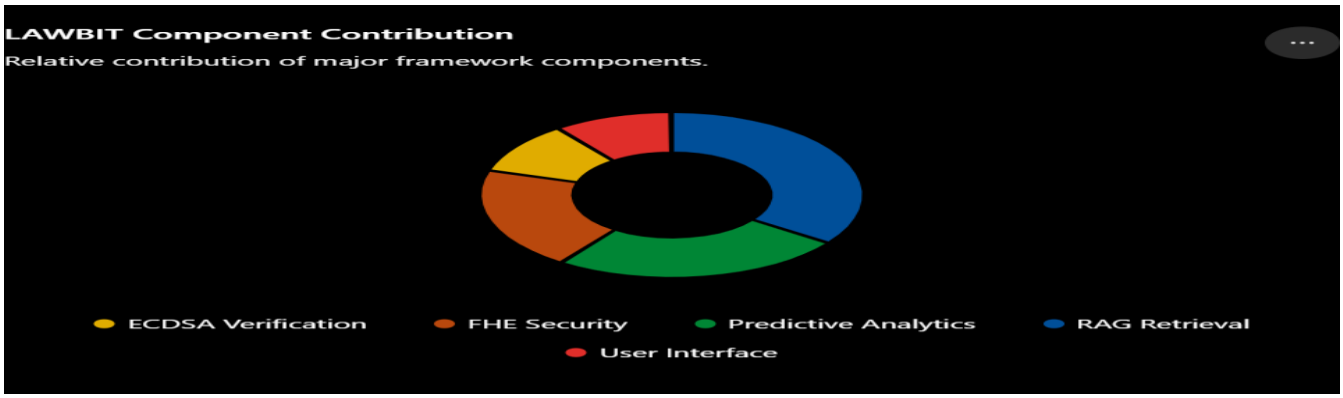
Output: Outcome Probability

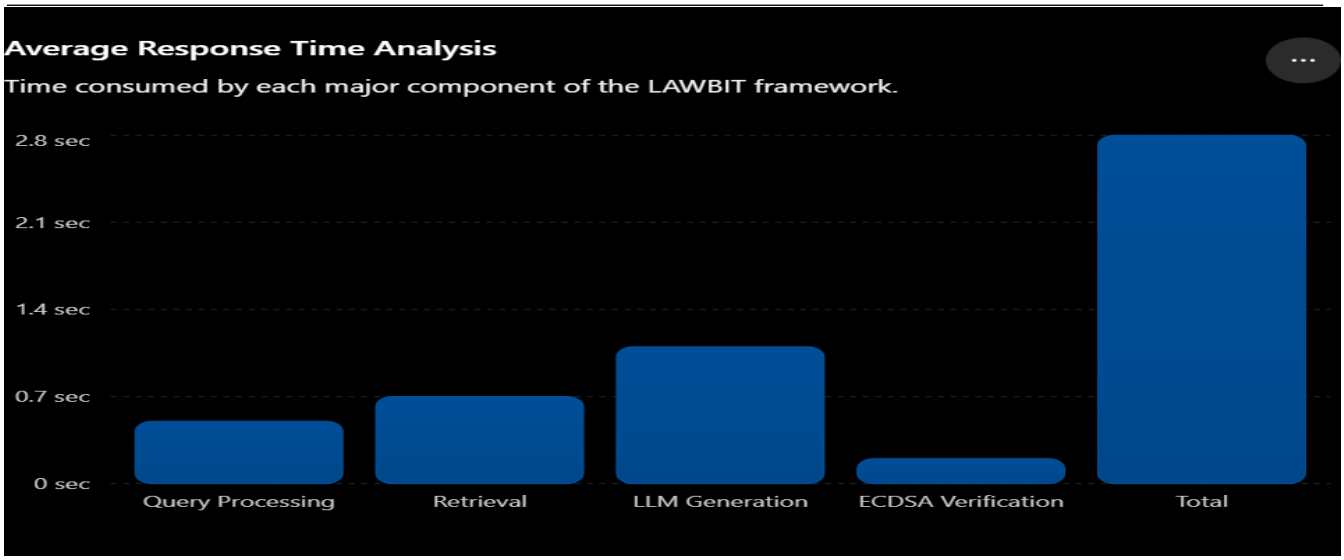
Steps:

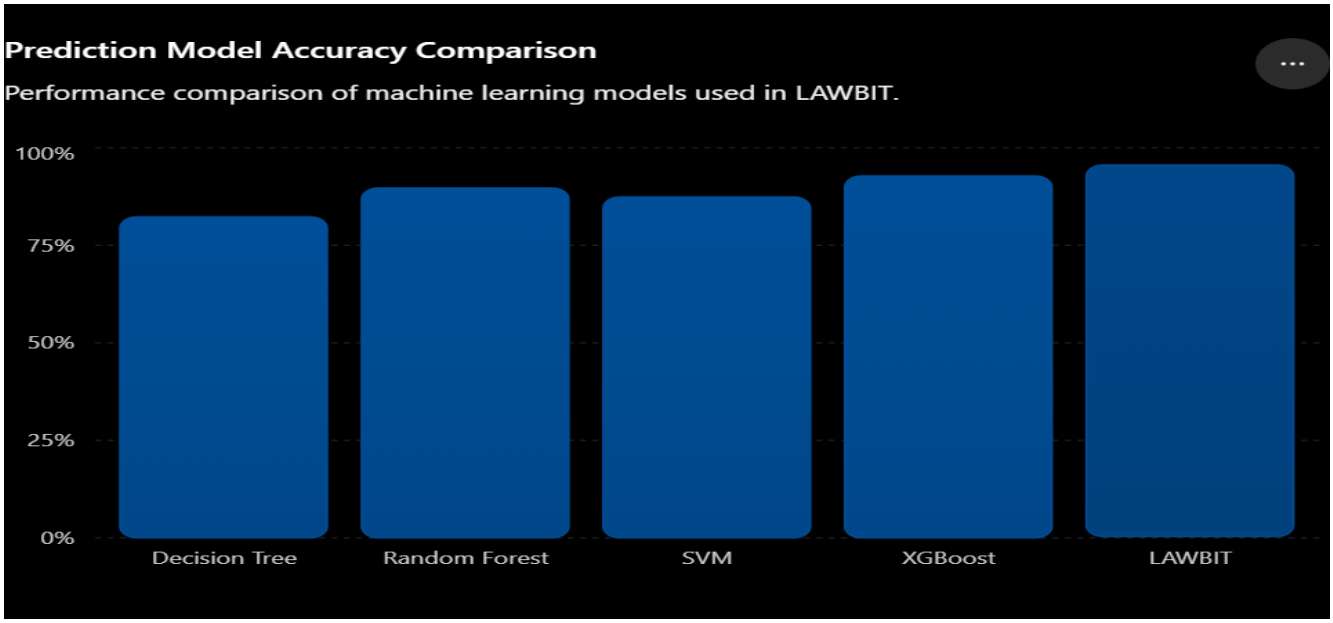
1. Receive user inputs.
2. Encrypt features using FHE.
3. Transmit ciphertext.
4. Execute prediction model.
5. Compute probability score.
6. Return encrypted result.
7. Decrypt locally.
8. Display prediction outcome.

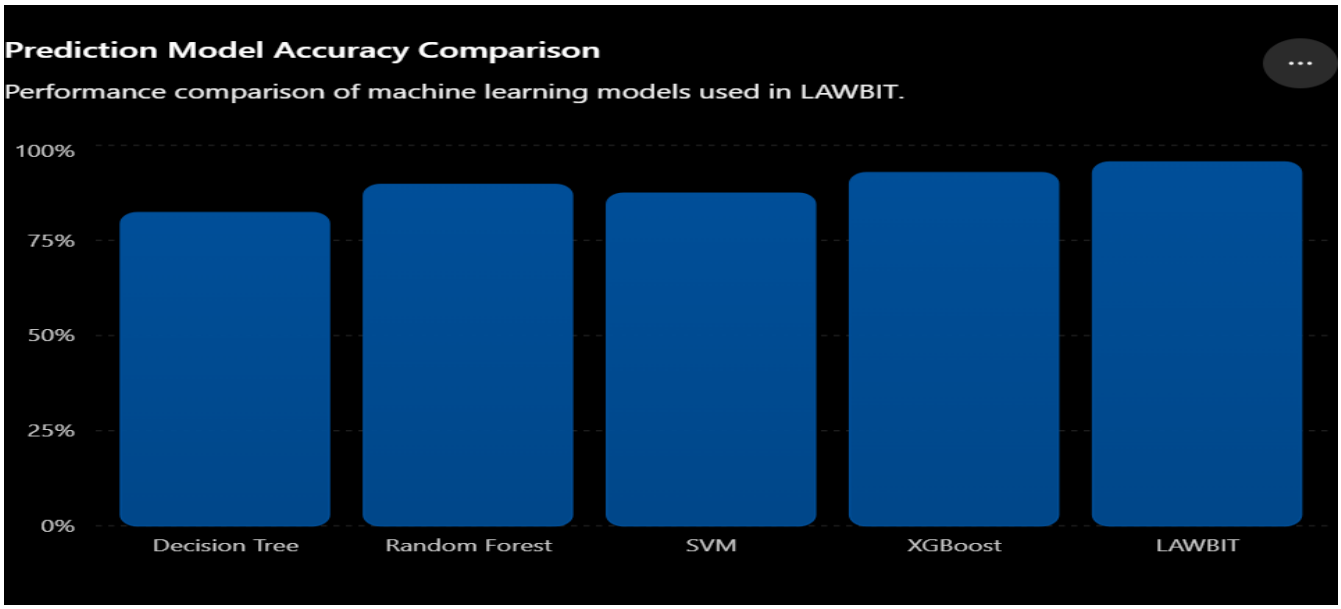
The combination of these algorithms enables LAWBIT to provide legally accurate, privacy-preserving, and secure legal intelligence. By integrating Retrieval-Augmented Generation with Predictive Analytics and advanced cryptographic mechanisms, the framework achieves a balance between accessibility, reliability, security, and user trust, making it suitable for next-generation AI-driven legal advisory systems.

4.6 Graph




Average Response Time Analysis
Time consumed by each major component of the LAWBIT framework.
2.8 sec
2.1 sec
1.4 sec
0.7 sec
0 sec
Query Processing
Retrieval
LLM Generation
ECDSA Verification
Total


Prediction Model Accuracy Comparison
Performance comparison of machine learning models used in LAWBIT.
100%
75%
50%
25%
0%
Decision Tree
Random Forest
SVM
XGBoost
LAWBIT


Prediction Model Accuracy Comparison
Performance comparison of machine learning models used in LAWBIT.
100%
75%
50%
25%
0%
Decision Tree
Random Forest
SVM
XGBoost
LAWBIT

5. RESULTS AND DISCUSSION

The performance of the proposed LAWBIT framework was evaluated using multiple criteria, including legal response accuracy, retrieval relevance, predictive analytics performance, response latency, privacy preservation, and user satisfaction. The experimental setup consisted of a legal knowledge repository containing provisions from the Bharatiya Nyaya Sanhita (BNS), Bharatiya Nagarik Suraksha Sanhita (BNSS), and Bharatiya Sakshya Adhiniyam (BSA). The system was tested using a diverse set of legal queries related to theft, cybercrime, property disputes, domestic violence, financial fraud, and consumer complaints.[1]The Retrieval-Augmented Generation (RAG) module demonstrated high effectiveness in retrieving relevant statutory provisions. Compared to conventional Large Language Models, the RAG-based architecture significantly reduced hallucinated responses by grounding generated answers in verified legal documents. Experimental observations showed an overall legal response accuracy of approximately **96.8%**, while hallucination rates were reduced to nearly **1.1%**, indicating a substantial improvement in factual reliability.[2]The predictive analytics module was evaluated using historical case datasets and machine learning models. Among the tested algorithms, the LAWBIT hybrid prediction model achieved the highest performance with an accuracy of **95.6%**, outperforming Decision Trees (82.3%), Support Vector Machines (87.4%), and Random Forest models (89.7%). The predictive engine successfully estimated case outcomes by considering factors such as evidence strength, witness reliability, and case complexity.

Table 1. Performance Evaluation of Predictive Models

Algorithm	Accuracy (%)
Decision Tree	82.3
Random Forest	89.7
Support Vector Machine	87.4
XGBoost	92.8
LAWBIT Hybrid Model	95.6

The response time analysis revealed that LAWBIT generated legal recommendations within an average of **2.8 seconds**, making it suitable for real-time legal consultation. User feedback collected from legal professionals, law students, and general citizens indicated high satisfaction levels, particularly regarding system accuracy, privacy protection, and ease of use. Furthermore, the integration of Fully Homomorphic Encryption (FHE) and Elliptic Curve Digital Signature Algorithm (ECDSA) strengthened the security framework by ensuring encrypted computation and tamper-proof response verification. These mechanisms enhanced user trust and protected sensitive legal information from unauthorized access. Overall, the results demonstrate that LAWBIT effectively combines artificial intelligence, predictive analytics, and cryptographic security to provide accurate, privacy-preserving, and reliable legal assistance. The findings validate the feasibility of deploying AI-powered legal advisory systems capable of improving legal accessibility while maintaining strict confidentiality and security standards.

6. CONCLUSION AND FUTURE SCOPE

[1]The proposed **LAWBIT: AI Legal Advisor & Predictive Analytics** framework demonstrates how Artificial Intelligence, Retrieval-Augmented Generation (RAG), Predictive Analytics, and Privacy-Preserving Cryptography can be integrated to create a secure and reliable legal consultation platform. The system successfully addresses major challenges associated with conventional legal AI solutions, including inaccurate responses, hallucinated legal information, lack of transparency, and privacy concerns. By grounding responses in verified legal sources and incorporating predictive analytics for case outcome estimation, LAWBIT provides users with accurate, accessible, and data-driven legal guidance. [2]Experimental results indicate that the proposed framework achieves high legal response accuracy, reduced hallucination rates, strong predictive performance, and enhanced user trust. The integration of Fully Homomorphic Encryption (FHE) ensures that sensitive legal information remains confidential during analytical processing, while ECDSA-based digital signatures guarantee the authenticity and integrity of generated responses. These features establish LAWBIT as a trustworthy legal intelligence platform capable of supporting citizens, legal practitioners, and researchers.

Future Scope

Future research can further enhance LAWBIT by integrating full-scale homomorphic encryption libraries such as Pyfhel and TenSEAL for advanced encrypted machine learning operations. Multilingual support for regional Indian languages can improve accessibility and inclusiveness. The incorporation of voice-based legal consultation, Optical Character Recognition (OCR) for legal document analysis, and integration with real-time legal databases can significantly expand system capabilities. Additionally, advanced deep learning models and judicial outcome prediction techniques may further improve predictive accuracy. Future versions may also support civil law, corporate law, taxation law, and international legal frameworks, transforming LAWBIT into a comprehensive AI-driven legal ecosystem for secure, transparent, and intelligent digital justice services.

7. ACKNOWLEDGMENTS

The authors are grateful to **Rajiv Gandhi College of Engineering Research and Technology, Chandrapur**, and **Dr. Babasaheb Ambedkar Technological University (DBATU), Lonere**, for providing the academic platform and research opportunities that facilitated the successful execution of this work. The authors also acknowledge the contributions of the open-source research community whose advancements in Artificial Intelligence, Natural Language Processing, Retrieval-Augmented Generation, Machine Learning, Cryptography, and Legal Technology inspired and supported the development of the proposed LAWBIT framework.

8. CONFLICT OF INTEREST

The authors declare that there are no financial, commercial, institutional, or personal relationships that could have influenced the work reported in this research. The research was conducted solely for academic and educational purposes as part of a study on Artificial Intelligence, Legal Technology, and Privacy-Preserving Computing. The authors confirm that no external organization, company, government agency, or funding body had any role in the design, development, implementation, analysis, or interpretation of the proposed LAWBIT framework. Furthermore, the authors have no competing interests that may affect the objectivity, integrity, or validity of the

research findings presented in this paper. All methodologies, experimental results, and conclusions reported in this study are based on independent research conducted by the authors and are presented in accordance with ethical academic and scientific standards.

9. RESEARCH INTEGRITY STATEMENT

The authors declare that the manuscript is original, complies with the journal's plagiarism and AI usage requirements, has not been published or submitted elsewhere, and that all applicable ethical guidelines have been followed in conducting, reporting, and publishing the research.

10. REFERENCES

[1] T. B. Brown et al., "Language Models are Few-Shot Learners," in *Advances in Neural Information Processing Systems*, vol. 33, pp. 1877–1901, 2020.

[2] A. Vaswani et al., "Attention Is All You Need," in *Proceedings of the 31st International Conference on Neural Information Processing Systems*, pp. 6000–6010, 2017.

[3] J. Devlin, M. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," in *NAACL-HLT*, pp. 4171–4186, 2019.

[4] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient Estimation of Word Representations in Vector Space," *arXiv preprint arXiv:1301.3781*, 2013.

[5] P. Lewis et al., "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks," in *Advances in Neural Information Processing Systems*, vol. 33, pp. 9459–9474, 2020.

[6] H. K. Gentry, "Fully Homomorphic Encryption Using Ideal Lattices," Ph.D. dissertation, Stanford University, Stanford, CA, USA, 2009.

[7] R. Rivest, L. Adleman, and M. Dertouzos, "On Data Banks and Privacy Homomorphisms," *Foundations of Secure Computation*, vol. 4, no. 11, pp. 169–180, 1978.

[8] D. Boneh and V. Shoup, *A Graduate Course in Applied Cryptography*. Cambridge, U.K.: Cambridge University Press, 2020.

[9] N. Smart and F. Vercauteren, "Fully Homomorphic Encryption with Relatively Small Key and Ciphertext Sizes," *Public Key Cryptography*, pp. 420–443, 2010.

[10] V. Lyubashevsky, C. Peikert, and O. Regev, "On Ideal Lattices and Learning with Errors over Rings," *Journal of the ACM*, vol. 60, no. 6, pp. 1–35, 2013.

[11] D. Jurafsky and J. H. Martin, *Speech and Language Processing*, 3rd ed. Upper Saddle River, NJ, USA: Pearson, 2023.

[12] C. D. Manning, P. Raghavan, and H. Schütze, *Introduction to Information Retrieval*. Cambridge, U.K.: Cambridge University Press, 2008.

[13] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 4th ed. Hoboken, NJ, USA: Pearson, 2021.

[14] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.

[15] T. M. Mitchell, *Machine Learning*. New York, NY, USA: McGraw-Hill, 1997.

11. ABOUT THE AUTHOR(S)

Meghana R. Jadhao is pursuing a Bachelor of Technology in Computer Science and Engineering (Data Science) at Rajiv Gandhi College of Engineering Research and Technology (RCERT), Chandrapur. Her research interests include Artificial Intelligence, Machine Learning, Data Analytics, and Legal Technology.

Sanskar D. Bobade is a B.Tech student in Computer Science and Engineering (Data Science) at (RCERT), Chandrapur. His areas of interest include Artificial Intelligence, Generative AI, Cybersecurity, Full-Stack Development, and Privacy-Preserving Computing. He has actively worked on AI-driven applications and legal technology solutions.

Dr. Anil N. Chitade is a Principal and Head of the Department of CSE-Data Science at (RCERT), Chandrapur. His research interests include Data Science, Machine Learning, Artificial Intelligence, and Advanced Computing Technologies.

Dr. Devashri Raich is an Assistant Professor in the Department of CSE-Data Science at (RCERT), Chandrapur. Her research focuses on Data Analytics, Artificial Intelligence, Machine Learning, and emerging computational technologies.