SMALL LANGUAGE MODELS FOR SRI LANKAN LEGAL APPLICATIONS

(Developing An Agent For Template Matching In Deed Document)

Project Final Report

E.S. Mathusigan – IT22177032

B.Sc. (Hons) in Information Technology Specializing in Information Technology

Department of Information Technology

Sri Lanka Institute of Information Technology

April 2026

SMALL LANGUAGE MODELS FOR SRI LANKAN LEGAL APPLICATIONS

(Developing An Agent For Template Matching In Deed Document)

Project Final Report

A. Thuvaraga– IT22030412

B.Sc. (Hons) in Information Technology Specializing in Information Technology

Department of Information Technology

Sri Lanka Institute of Information Technology

April 2026

DECLARATION

I declare that this is our own work, and this proposal does not incorporate without acknowledgment any material previously submitted for a degree or diploma in any other university or Institute of higher learning, and to the best of our knowledge and belief, it does not contain any material previously published or written by another person except where the acknowledgment is made in the text. Also, I hereby grant to the Sri Lanka Institute of Information Technology, the nonexclusive right to reproduce and distribute my dissertation, in whole or in part in print, electronic or other mediums. I retain the right to use this content in whole or part in future works (such as articles or books).

Name	Student ID	Signature
A. Thuvaraga	**IT22030412**	

The supervisor/s should certify the proposal report with the following declaration.

The above candidates are carrying out research for the undergraduate Dissertation under my Supervision.

Signature of the co -supervisor

Signature of the supervisor

Date

Date

ABSTRACT

Deeds are the pillars of transfer of property and inheritance and should be proper and consistent to prevent litigation. Human verification is time-consuming and expensive and liable to miss details in bulky papers or poor scans. Extremely sophisticated AI systems can reduce it to some extent but are too expensive and complicated with restricted resources. We investigate a possible alternative in this research work: small language models (SLMs) with explicit rule-based template matching to identify anomalies in deeds. We process with five common document classes Power of Attorney, Deed of Transfer/Conveyance (Sale Deed), Deed of Gift, Last Will & Testamentary Deed, and Deed of Mortgage and produce two outputs automatically: (1) a concise document summary and (2) a conformity report that highlights each alert with the exact lines where issues exist. We have a four-step pipeline. Step 1: Ingest and validation. We ingest the file, verify format and size and if scanning, conduct OCR and record a quality score. Step 2: Information extraction. Light-weight models and pattern rules are trained on legal texts and pull out major fields (parties, dates, property description, amounts, witnesses, notary details, encumbrances).Step 3: Deed classification and agent routing. A deed-type router classifies and sends the document to the appropriate deed agent. Within each agent, we run a template-matching algorithm that integrates (a) firm rules for mandatory clauses and cross-checks (e.g., attestation, chain of title, acceptance of gift, executor appointment, loan and security terms) with (b) narrow SLM prompts for intent and language consistency (e.g., excessively broad powers, latent conditions, unclear boundaries to property). The algorithm produces issue type, severity, and evidence ranges (page/line). Step 4: Scoring and explanation. Checks are given a score of Low/Medium/High and tabulated as a total conformity score. A confidence measure is calculated from OCR quality and extraction coverage. We tested the method on a mix of clean PDFs and noisy scans. Routing to the right agent plus the agent-internal template checks helped small models work well with narrow prompts and deed-specific rules. The system consistently flagged important problems such as missing attestation details, name mismatches, unclear boundaries, absent acceptance language in gift deeds, overwide powers in powers of attorney, and incomplete release terms in mortgages. Reviewers reported that evidence-linked findings made verification faster and more reliable. The system ran quickly and showed a clear “low confidence” warning when text quality was poor.

Keywords— Small Language Model (SLM) ,Agentic Retrieval-Augmented Generation ,Agents, Natural Language Processing (NLP), Transformer-based Models, Qwen, Retrieval-Augmented Generation (RAG), FAISS, PostgreSQL, Semantic Similarity Search, Domain-Specific Language Models, Sri Lanka Legal System.

ACKNOWLEDGEMENT

The success of this research would not have been possible without the support and guidance of many individuals.

First, I would like to express my sincere gratitude to my supervisor, Dr. Prasana Sumathpala, and co-supervisor, Ms. Karthiga Rajendran, for their continuous guidance, encouragement, and valuable advice throughout the research project. Their expertise and constructive feedback were essential in completing this work successfully.

I would also like to extend my thanks to all lecturers, instructors, assistant lecturers, and both academic and non-academic staff of the Sri Lanka Institute of Information Technology (SLIIT) for their support and contributions during this journey.

My heartfelt appreciation goes to my group members for their cooperation, teamwork, and assistance during the research process.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Abbreviations	Descriptions
ML	Machine Learning
NLP	Natural Language Processing
SLM	Small Language Model
RAG	Retrieval-Augmented Generation
NER	Named Entity Recognition
API	Application Programming Interface
AI	Artificial Intelligence
OCR	Optical character recognition
QLoRA	Quantize Low-Rank Adaptation

1. INTRODUCTION

The increasing complexity of legal processes and the growing volume of documentation in Sri Lanka have created significant challenges in ensuring accuracy, efficiency, and accessibility within the legal system. Legal documents such as deeds play a fundamental role in governing property transfer, inheritance, and contractual obligations. Errors, omissions, or ambiguities in such documents can lead to costly disputes, prolonged litigation, and administrative inefficiencies. Traditionally, the verification and interpretation of these documents rely heavily on manual review by legal professionals, which is time-consuming, expensive, and susceptible to human error—particularly when dealing with lengthy documents or low-quality scanned copies.

Recent advancements in Artificial Intelligence (AI), particularly in Natural Language Processing (NLP), have introduced opportunities to automate aspects of legal analysis. However, most state-of-the-art solutions rely on large language models (LLMs), which demand substantial computational resources, specialized infrastructure, and high operational costs. These requirements limit their practical applicability in resource-constrained environments such as Sri Lanka. Furthermore, many of these systems lack transparency, making it difficult for legal practitioners to trust and validate their outputs.

To address these challenges, this research proposes a comprehensive, cost-effective legal technology framework tailored to the Sri Lankan context. The system is designed to deliver real-world applicability by leveraging small language models (SLMs) combined with rule-based methodologies, ensuring both efficiency and explainability. The framework is built upon four key pillars that collectively aim to enhance legal accessibility, accuracy, and decision-making.

The first pillar focuses on the fine-tuning of small language models to provide end-to-end, step-by-step legal guidance in specific domains, particularly Sri Lankan real estate law and family law. By training SLMs on curated legal texts, statutes, and domain-specific materials, the system is capable of generating structured and context-aware legal instructions. This enables users, including non-experts, to better understand legal procedures and requirements without the need for extensive legal consultation.

The second pillar, which forms the core component of this research, is a template-based deed verification system. This component integrates lightweight SLMs with explicit rule-based checks to automate the analysis of legal deeds. The system is designed to process five commonly used document types in Sri Lanka: Power of Attorney, Deed of Transfer (Sale Deed), Deed of Gift, Last Will and Testamentary Deed, and Deed of Mortgage. Upon ingestion of a document in either PDF or scanned format, the system performs optical character recognition (OCR) when necessary and extracts key information such as parties involved, dates, property descriptions, financial details, witnesses, and notarial elements.

Following information extraction, the system employs a classification mechanism to identify the type of deed and routes it to a specialized, deed-specific agent. Each agent applies a combination of predefined template rules and targeted SLM prompts. The rule-based component ensures the presence of mandatory clauses and validates structural consistency, including attestation requirements, chain of title, acceptance clauses, and financial obligations. Meanwhile, the SLM component focuses on detecting semantic inconsistencies, such as ambiguous language, overly broad authorizations, or unclear property boundaries.

The output of this process consists of two key deliverables: a concise document summary and a detailed conformity report. The conformity report highlights deviations from expected template structures, categorizes issues by severity (e.g., low, medium, high), and provides precise page and line references as evidence. Additionally, the system generates a confidence score based on factors such as OCR quality and completeness of extracted data, enabling users to assess the reliability of the analysis.

The third pillar of the framework introduces a law recommendation system aimed at assisting citizens, with an initial focus on labor and employment law. This component is designed to provide accessible legal guidance based on user queries, helping individuals understand their rights and obligations in common legal scenarios.

The fourth pillar involves the development of predictive models capable of estimating probable legal outcomes based on historical case data. By analyzing past judgments, these models aim to support informed decision-making and risk assessment for both legal practitioners and individuals.

The proposed system was evaluated using a combination of clean digital documents and noisy scanned inputs. Experimental results demonstrate that the integration of template-based verification with small language models effectively identifies common legal errors, including missing attestation details, inconsistencies in names, incomplete clauses, and ambiguous wording. Furthermore, the system significantly reduces the time and cost associated with manual review while maintaining a high level of transparency and auditability.

In summary, this research presents a practical and scalable legal AI solution tailored to the Sri Lankan legal ecosystem. By combining efficiency, interpretability, and domain-specific design, the proposed framework offers a viable alternative to resource-intensive AI systems, contributing to the advancement of accessible and reliable legal technology.

1.1 Background and Literature Survey

Property succession and transfer in Sri Lanka are fundamentally governed by legal documents such as Powers of Attorney, Sale Deeds, Gift Deeds, Testamentary Deeds, and Mortgage Deeds. These documents must be precise, legally valid, and internally consistent to ensure the legitimacy of transactions and to prevent disputes. In practice, however, most deeds are lengthy, formally structured, and often exist as stamped, signed, scanned, or printed copies. This makes manual verification a labor-intensive and time-consuming process. Furthermore, human reviewers may overlook critical inconsistencies, especially when dealing with large volumes of documents or poor-quality scans.

Recent advancements in Artificial Intelligence (AI), particularly large language models (LLMs), have demonstrated the potential to automate legal document analysis. These systems can assist in reviewing and interpreting complex legal text. However, their practical adoption in contexts such as Sri Lanka is limited due to high computational costs, infrastructure requirements, and concerns related to data privacy when handling sensitive land and property records. As a result, there is a growing need for more efficient, cost-effective, and privacy-aware solutions.

A promising alternative is the use of small language models (SLMs) combined with transparent, rule-based template matching approaches. Unlike large models, SLMs can operate with lower computational requirements while still providing meaningful contextual understanding when applied to narrow, domain-specific tasks. By integrating SLMs with explicit template rules, it becomes possible to create systems that not only detect anomalies in legal documents but also provide clear, evidence-based explanations for each finding. This approach aligns well with the requirements of legal verification, where interpretability and traceability are critical.

The primary objective of this research is to develop a “deed verification agent” capable of automatically identifying the type of deed, extracting key information, enforcing domain-specific rules, and applying focused SLM-based reasoning. The system is designed to generate a concise summary, a conformity score, and detailed evidence in the form of page and line references. Such a system aims to support legal professionals by reducing manual workload while improving accuracy and consistency.

Existing literature provides a strong foundation for this approach. Prior studies have demonstrated that machine learning techniques can effectively identify irregular or anomalous text in legal agreements and provide simplified signals to assist human reviewers [1], [24]. More recent research has focused on clause-level template matching, where models are trained to detect missing, incomplete, or improperly structured clauses. These methods are particularly relevant for deed verification, as they enable systematic validation of document structure and content [2].

At the system level, several studies advocate for hybrid approaches that combine expert-defined rules with retrieval-augmented generation (RAG) and iterative refinement techniques. These approaches have been shown to reduce errors and improve trust in AI-generated outputs [3], [4], [22]. In the context of deed verification, such methods are especially valuable, as they allow the system to reference specific portions of a document and, when necessary, link findings to relevant legal statutes or practice guidelines.

Another important area of research focuses on processing long and complex legal documents. Studies have shown that segmenting large documents into smaller sections improves model performance and enables more accurate analysis [5]. Additionally, neural network-based approaches have demonstrated superior performance compared to traditional methods in understanding legal text [6]. Despite these advancements, challenges remain in handling very long documents, supporting multiple languages, and ensuring that outputs are both accurate and explainable [7], [8].

In the domain of summarization, recent work emphasizes the importance of precise, citation-based summaries rather than generic abstractions. This is particularly relevant for legal applications, where summaries must retain key details and provide verifiable references [9]. Such approaches are directly applicable to deed verification systems, which require concise yet evidence-backed outputs.

Although judgment prediction research does not directly target deed analysis, several techniques from this domain are highly relevant. Multi-task learning approaches, which jointly model related tasks, have been shown to improve accuracy and robustness [10]. Similarly, feedback-driven methodologies help reduce errors, while semantic extraction techniques enable the identification of critical facts within documents [11], [12]. These ideas can be adapted to deed verification, particularly for cross-checking information such as names, dates, and property details across different sections of a document.

Closer to the domain of contracts and deeds, research has demonstrated that incorporating legal domain knowledge into language models significantly enhances clause detection and validation capabilities [13]. Specialized models such as LEGAL-BERT have established strong baselines for clause-level analysis in legal texts [14], [17]. Additionally, datasets such as ACORD support clause ranking and retrieval tasks, which are essential for identifying specific text segments that comply with or violate template rules [15], [16]. Evaluation frameworks like LegalBench and RAG benchmarking tools further provide standardized methods for assessing retrieval and reasoning performance in legal AI systems [22].

In the broader context of document verification, there is a strong emphasis on explainability and uncertainty estimation. AI systems must not only provide results but also clearly explain their reasoning and indicate confidence levels to ensure user trust [23]. Recent studies highlight that small, efficient models combined with rule-based systems and retrieval mechanisms can deliver reliable performance in low-resource and privacy-sensitive environments [20]. Furthermore, systems that ground their outputs in actual laws and regulations are more trustworthy and applicable in real-world legal settings, particularly in jurisdictions such as Sri Lanka [21], [4].

In summary, the existing body of research supports the development of a hybrid, explainable, and resource-efficient approach to legal document verification. By integrating small language models with template-based rule systems, it is possible to create a practical solution tailored to the needs of Sri Lankan legal practice, addressing both technical and operational challenges.

Table 1:LLM Used Papers

LLM-Used Papers

Paper	**Technologies Used**	**Strengths**	**Limitations**
[2] Contract Eval	LLM prompts; clauselevel risk labels; evaluation suite	Standardized benchmark; focuses on risk at clause level; useful metrics	Contracts domain (not deeds); largemodel cost; limited multilingual
[3] Reliable Legal AI (modular)	Expert rules RAG; guardrails/refinement; pipeline orchestration	Reduces hallucinations; auditability; explainable workflow	Complex to build; needs curated KB; not deedspecific
[4]/[22] Bridging Legal Knowledge & AI	Vector store; knowledge graph; hierarchical NMF; retrieval + grounding	Strong grounding; better snippet retrieval; interpretable links	KB creation/maintenance overhead; latency; infra heavy
[7] Legal NLP Survey (2024)	Survey of LLM tasks/datasets/models	Broad landscape view; gaps & challenges summarized	No system or code; highlevel only
[8] Legal NLP Survey (2023)	Survey across pre-LLM and LLM era	Historical context; taxonomy of tasks	Not implementation - focused; jurisdictionagnostic

[9] Legal Summarization Survey (2025)	LLM/extractive–abstractiv e summarizers; factuality checks	Guidance for evidencelinked summaries; eval pitfalls	Survey only; no ready model
[13] Knowledge-Augmented LLM for Contract Risk	Domain knowled ge injection; prompted LLM risk flags	Better risk detection with domain hints; transferable idea	Construction contracts focus; knowledge curation effort; compute cost
[15]/[16] ACORD (Clause Retrieval Dataset)	Biencoder retrieval; crossencoder/ LLM reranking	Large expert dataset; trains evidence retrieval	Contract clauses (not deeds); mainly English; setup effort
[21] LawGPT (legal LLM, CN)	Legal corpus pretrainin g; instructio n tuning; knowled geenhanced	Domain gains for QA/reasoning; template for legal tuning	Non-Sri Lankan law; big model footprint; adaptation required

The studies summarized in Table 2 highlight the growing role of large language models (LLMs) and hybrid AI systems in legal document analysis, while also revealing important limitations that motivate this research. Works such as [2] (*Contract Eval*) demonstrate the effectiveness of LLM prompts for clause-level risk identification and provide standardized evaluation benchmarks. However, these approaches are primarily designed for contracts rather than deeds and rely on large models, making them costly and less suitable for low-resource environments.

Research on modular legal AI systems, such as [3], emphasizes combining expert rules with retrieval-augmented generation (RAG) and guardrails to improve reliability and reduce hallucinations. While these systems enhance explainability and auditability, they are complex to implement and require curated knowledge bases. Similarly, studies like [4] and [22] focus on integrating knowledge graphs and vector-based retrieval to improve grounding and interpretability. Although effective, these methods introduce additional infrastructure complexity and maintenance overhead.

Survey papers such as [7] and [8] provide a broad overview of legal NLP advancements, identifying key challenges such as handling long documents, multilingual support, and ensuring trustworthy outputs. In the area of summarization, [9] highlights the importance of evidence-linked summaries, which directly aligns with the needs of deed verification systems.

Other works, including [13] and datasets like [15]/[16] (ACORD), demonstrate that domain knowledge integration and clause retrieval significantly improve risk detection. However, these are largely focused on contract domains and require substantial setup effort. Finally, systems like [21] (*LawGPT*) show the benefits of domain-specific training but are not tailored to Sri Lankan law and involve large computational footprints.

Overall, these studies support the need for a lightweight, explainable, and domain-specific approach, reinforcing the relevance of using small language models combined with rule-based verification for deed analysis.

Agents-Used Papers

Paper	Technologies Used	Strengths	Limitations

[3] Reliable Legal AI (modular)	Expert rules; RAG (vector store/KG); guardrails/refinem ent; orchestration	Lowers hallucinations; explainable and auditable; modular, domain-grounded	High integration effort; curated KB needed; latency/ops overhead; not deedspecific
[4]/[22] Bridging Legal Knowled ge & AI	Retrieval stack: vector store + KG + hierarchical NMF; RAG pipeline	Precise snippet retrieval; interpretable links; stronger grounding	KB/KG build & maintenance cost; heavier infra; multilingual setup needed
[20] SLMs for Agentic AI	Small LMs; tooluse; planners; multi-agent orchestration; quantization	Low cost; fast; onprem friendly; composable agents	Narrower knowledge than big LLMs; needs careful task decompositi on; more engineering glue
[11] MultiPerspecti ve Bi-Feedback Network	Predict-then-check loop; attention; multi-signal training	Built-in verification reduces errors; robust pattern for 'verify' stage	Requires labeled data; not a deed pipeline; explanations limited without spans

[10] Topological MultiTask LJP	DAG/graph of dependent subtasks; joint training	Captures task dependencies;improves consistency; informs agent step ordering	Case/judgment focus (not deeds); needs welldefined task labels; transfer requires mapping

Table 2:Agents-Used Papers

The studies summarized in Table 3 highlight the evolution of agent-based approaches in legal AI systems, focusing on modularity, reliability, and efficiency. Research such as [3] (*Reliable Legal AI*) demonstrates the effectiveness of combining expert rules, retrieval-augmented generation (RAG), and guardrails within an orchestrated pipeline. This approach significantly reduces hallucinations and improves explainability and auditability. However, it requires substantial integration effort, curated knowledge bases, and introduces operational overhead, making it less suitable for lightweight applications.

Similarly, works like [4] and [22] emphasize advanced retrieval architectures that integrate vector stores, knowledge graphs, and hierarchical methods to improve grounding and snippet-level accuracy. While these systems enhance interpretability and precision, they demand significant infrastructure and maintenance, especially when adapting to multilingual legal environments.

In contrast, [20] explores the use of small language models (SLMs) within agentic frameworks, highlighting their advantages in cost efficiency, speed, and suitability for on-premise deployment. These systems are highly modular but require careful task decomposition and additional engineering to coordinate multiple agents effectively.

Other approaches, such as the bi-feedback verification model in [11], introduce a “predict-then-check” mechanism that improves reliability through internal validation loops. Additionally, [10] demonstrates how multi-task learning with dependency graphs can enhance consistency across related tasks.

Overall, these studies support the adoption of modular, agent-based architectures, particularly those leveraging SLMs, for building efficient and explainable deed verification systems.

Paper	Technologies Used	Strengths	Limitations
[1] Chakrabarti et al. (2019)	Doc2Vec; traditional classifiers (SVM/LogReg); paragraph-level risk scoring	Compute-light; easy to train; explainable features	Limited context understanding; manual features; not deed-specific
[5] Long-length Legal Doc Classification (2019)	Segmenting long docs; BiLSTM/attention; chunk aggregation	Works with long PDFs; small models; simple deployment	Loses crosssegment context; requires careful chunking
[6] Empirical DL for Legal Review (2018)	CNN/compact neural models vs classical baselines	Better than SVMs with enough data; fast inference	Needs labeled data; weaker on nuanced reasoning
[10] Topological LJP (2018)	DAG of dependent subtasks; multitask learning (non-LLM)	Captures task dependencies; efficient	Case-law focus; mapping to deeds required
[11] Bi-Feedback LJP (2019)	Predict-then-check loop; attention; compact models	Built-in verification; reduces false positives	Needs curated labels; explanations limited without spans

[12] SLJP (2023)	Transformer encoder; semantic extraction; chunked processing	Strong accuracy with moderate size; good for long text	Not pretrained on Sri Lankan deeds; GPU helpful for training
[14] Leveraging BERT for Legal Classification (2021)	BERT-base; domain tuning; long-doc strategies	Solid baseline; adaptable to deeds; moderate compute	512-token limit; needs segment/merge logic
[17] LEGAL-BERT (2020)	Legal-domain BERT variants; domain pretraining	Better legal vocabulary; drop-in backbone for NER/classification	Englishcentric; still mid-size; local corpus needed
[18] RealEstate Risk Assessment (2021)	ANP/MCDA scoring; lightweight analytics	Clear weighting of factors; inspires risk aggregation	Non-NLP; no text understanding; domain shift to deeds
[19] XAI for Credit Risk (2025)	Tree/linear models; SHAP/LIME explanations	Transparent scoring; user-friendly justifications	Tabular focus; needs adaptation for text spans
[24] AI in Legal Domain (2019)	Classic ML/NLP pipeline overview	Historical baselines; low compute	Outdated vs transformers; high manual feature work

Table 4:Small Language Model–Used Papers

The studies summarized in Table 4 highlight the evolution of lightweight and small-model approaches for legal document analysis, emphasizing efficiency, explainability, and practicality. Early works such as Chakrabarti et al. [1] demonstrate the use of Doc2Vec and traditional classifiers like SVM and Logistic Regression, which are computationally efficient and easy to train. However, these methods lack deep contextual understanding and rely heavily on manual feature engineering.

Research on long legal document processing, such as [5], introduces segmentation techniques combined with BiLSTM and attention mechanisms to handle large texts. While effective for long PDFs, these approaches often lose cross-section context. Similarly, compact deep learning models like CNNs [6] improve performance over classical methods but require labeled datasets and struggle with nuanced reasoning.

Transformer-based approaches, including BERT [14] and LEGAL-BERT [17], provide stronger contextual understanding and domain adaptation, making them suitable baselines for legal tasks. However, they are limited by token constraints and often require document chunking strategies. More recent works like [12] combine transformer encoders with semantic extraction to improve performance on long texts, though domain-specific training is still required.

Additionally, non-NLP methods such as risk scoring models [18] and explainable AI techniques like SHAP/LIME [19] contribute to transparency but lack direct text understanding. Overall, these studies support the use of efficient, moderately sized models combined with explainability techniques, aligning well with SLM-based deed verification systems.

1.1.1 Proposed solution

This research proposes a **hybrid, agent-based deed verification system** tailored for the Sri Lankan legal context, combining small language models (SLMs) with rule-based template matching to achieve accurate, explainable, and cost-efficient document analysis. The system is designed to process five common deed types—Power of Attorney, Sale Deed, Gift Deed, Last Will, and Mortgage Deed—commonly used in property transfer and succession.

The solution follows a structured pipeline. First, documents are ingested in PDF or scanned format, with Optical Character Recognition (OCR) applied when necessary, along with a quality assessment. Next, key information such as parties, dates, property details, and financial terms is extracted using lightweight models and pattern-based rules. A classification module then identifies the deed type and routes the document to a specialized agent.

Each agent applies predefined legal templates to verify mandatory clauses and structural consistency, while SLM-based prompts analyze semantic issues such as ambiguous language or missing intent. Detected issues are categorized by severity and linked to exact page and line references for transparency.

The system produces a concise summary, a conformity score, and a confidence indicator based on data quality. This approach ensures faster, more reliable, and explainable deed verification while remaining practical for low-resource environments.

1.2 Research Gap

Despite the rapid advancement of Artificial Intelligence (AI) in legal document analysis, the application of these technologies to Sri Lankan deed verification remains significantly underexplored. Existing research and commercial tools predominantly focus on contracts, case law, or legal systems in developed jurisdictions, leaving a critical gap in addressing the unique requirements of Sri Lankan property and inheritance documents. This research identifies several key gaps that limit the effectiveness, applicability, and adoption of AI-driven solutions in this domain.

A fundamental limitation is the absence of publicly available, well-annotated datasets specific to Sri Lankan deeds. Unlike other legal domains where benchmark datasets exist, there is no standardized corpus covering major deed categories such as transfers, gifts, powers of attorney, mortgages, and testamentary documents. This restricts the ability to train, evaluate, and compare machine learning models in a consistent and reproducible manner.

Another critical gap lies in the lack of machine-enforceable local legal rules. Sri Lankan deeds follow specific legal and procedural requirements, including attestation details, acceptance clauses in gift deeds, chain of title verification, and priority or release conditions in mortgage documents. However, these domain-specific rules are not formally encoded into computational frameworks, limiting the ability of AI systems to perform accurate and context-aware validation.

Real-world deed processing is further complicated by the poor quality and variability of document formats. Many deeds exist as low-resolution scans with faded text, handwritten annotations, skewed layouts, and embedded elements such as tables or survey plans. Current research systems, which are often evaluated on clean and structured datasets, do not adequately address these challenges, resulting in reduced reliability in practical settings.

A significant usability limitation is the lack of evidence-linked outputs. Legal professionals require precise references—such as page and line numbers—to verify and trust system-generated findings. However, most existing AI solutions provide only high-level predictions or scores without linking them to specific text spans, thereby limiting transparency and auditability.

Additionally, much of the current work relies on large, cloud-based language models, which are computationally expensive and raise concerns regarding data privacy and regulatory compliance. In contrast, the potential of small language models (SLMs) operating with rule-based template matching in secure, on-premises environments remains largely unexplored.

There is also a notable lack of legal grounding in outputs, as many systems fail to reference relevant Sri Lankan statutes, circulars, or professional practice guidelines. This reduces the practical value and trustworthiness of AI-assisted legal analysis. Furthermore, the absence of a standardized evaluation framework, including issue taxonomies, severity classifications, and annotated benchmarks, makes it difficult to assess system performance objectively.

Temporal factors introduce another layer of complexity. Legal requirements may vary over time, and most systems do not account for the applicable law at the date of deed execution, leading to potential inaccuracies. Moreover, existing approaches rarely incorporate human-in-the-loop feedback mechanisms, such as corrections from notaries or conveyancers, which are essential for continuous system improvement.

From a system design perspective, there is a clear gap in end-to-end workflow integration. Most research focuses on isolated tasks such as OCR, named entity recognition, or classification, whereas real-world applications require a complete pipeline—from document ingestion and validation to classification, extraction, verification, reasoning, legal retrieval, scoring, and reporting.

Finally, practical deployment considerations remain insufficiently addressed. Sri Lankan legal environments often require privacy-preserving, on-premises solutions with secure storage and audit capabilities, which are not commonly supported by existing systems. In addition, fairness and robustness across languages and document variations—including Sinhala, Tamil, and older deed formats—are rarely evaluated, potentially leading to biased or inconsistent performance.

In summary, these gaps highlight the need for a domain-specific, explainable, resource-efficient, and end-to-end legal AI system tailored to Sri Lankan deed verification, which this research aims to address.

Deed template matching Tools Comparison

Product	Primary scope	Key tech / features	Strengths	Limitations vs your goals	Fit for Sri Lankan deed risk
Deed Reader Pro	Metes-andbounds plotting from deed text	OCR → parse bearings/distances → CAD/IntelliCAD export; Windows app; free trial	Very fast plotting; surveyorfriendly outputs; practical desktop tool	No legal risk analysis; no statute grounding; not Sri Lanka-specific; geometry focus	Low
Deed Plotter AI	AI plotting for deeds/easements/leases	OCR + NLP/LLM → JSON → map/plot; high-volume processing	Automat es plotting at scale; integratio ns; modern web tool	No clause/issue risk checks; not deed-type agents; not localized to SL law	Low
V7 Go – Deed Analysis Agent	Title-exam extraction & summarization	Agentic workflow; reads deeds; extracts ownership data, legal descriptions, encumbrances	Speeds title review; agent platform; enterpris e support	Black-box LLM stack; cloud/infra heavy; not tailored to SL deed rules; unclear evidence-span granularity	Medium

AiPazz	Legal research (Sri Lanka)	Search Sri Lankan cases/legislation; AI insights	Local database; helpful for research and citations	Not a deed parser or risk engine; no OCR/plotting	Medium (as a legal source, not analyzer)
Your SLM deed-risk agent	Risk analysis for 5 deed types (PoA, Transfer/Sale, Gift, Testamentary, Mortgage)	On-prem SLM + rules + optional RAG; OCR (Si/Ta/En); deed classifier; issue severity + evidence spans + confidence	Evidence -linked, explaina ble; Sri Lanka rule checklist s; privacy-first; lightweig ht	Needs local dataset/annotations; OCR robustness; future registry/RAG links	High

1.2.1 Critical Analysis of Existing Research

Existing research in legal AI demonstrates significant progress in automating document analysis, yet it reveals several limitations when applied to the domain of Sri Lankan deed verification. A large portion of current work focuses on contracts, litigation documents, or jurisdictions outside Sri Lanka, resulting in solutions that are not directly transferable to local legal practices. While these systems perform well in structured and resource-rich environments, their applicability in low-resource, domain-specific contexts remains limited.

Many advanced approaches rely on large language models (LLMs), which offer strong capabilities in understanding and generating legal text. However, these models introduce challenges related to computational cost, infrastructure requirements, and data privacy. In legal domains—particularly those involving sensitive property records—cloud-based processing raises concerns about confidentiality and regulatory compliance. Additionally, LLMs often lack transparency, making it difficult for legal professionals to interpret and trust their outputs without clear supporting evidence.

Another key limitation is the lack of domain-specific rule integration. Although some studies incorporate retrieval-augmented generation (RAG) or knowledge graphs to improve grounding, they often require complex system architectures and extensive knowledge base maintenance. Furthermore, most research does not adequately address the need for machine-enforceable legal templates, which are essential for verifying structured documents such as deeds.

Technical challenges also persist in handling long and unstructured documents. While segmentation and transformer-based models improve performance, they may lose cross-sectional context or require additional processing logic. Moreover, many systems provide high-level predictions without linking results to exact text spans, reducing explainability and auditability.

Finally, existing research often lacks end-to-end integration, focusing on isolated tasks rather than complete workflows. Issues such as multilingual support, temporal legal variations, and real-world deployment constraints are insufficiently explored. These limitations highlight the need for a lightweight, explainable, and domain-adapted solution, reinforcing the relevance of the proposed hybrid SLM-based approach.

1.2.2 Critical Evaluation of Existing Systems

Existing systems in legal AI have demonstrated notable progress in automating document understanding, classification, and analysis. However, a critical evaluation reveals that most of these systems are not fully aligned with the practical requirements of deed verification, particularly within the Sri Lankan legal context.

A significant portion of current systems is built on large language models (LLMs), which provide strong capabilities in semantic understanding and text generation. While these systems perform well in tasks such as contract analysis and legal question answering, they are often computationally expensive and dependent on cloud-based infrastructure. This limits their feasibility in environments where cost, privacy, and data security are major concerns. Furthermore, many LLM-based systems lack transparency, making it difficult for legal practitioners to verify how conclusions are derived.

Another limitation is the heavy reliance on generalized legal datasets. Most systems are trained on contracts or case law from foreign jurisdictions, which differ significantly in structure and legal requirements from Sri Lankan deeds. As a result, these models struggle to capture domain-specific nuances such as attestation requirements, property boundary descriptions, and deed-specific clauses.

Although some systems incorporate advanced techniques such as retrieval-augmented generation (RAG) and knowledge graphs to improve grounding, they often introduce additional complexity in terms of system design, maintenance, and computational overhead. These approaches also require well-curated knowledge bases, which are not readily available for Sri Lankan legal materials.

From a technical perspective, handling long and unstructured documents remains a challenge. While segmentation and transformer-based methods improve performance, they may lose contextual relationships across sections. Additionally, most systems focus on isolated tasks—such as OCR, entity recognition, or classification—without providing a fully integrated, end-to-end workflow required for real-world applications.

A critical shortcoming across many systems is the lack of explainability. Outputs are typically presented as labels or risk scores without linking them to specific text evidence, reducing user trust and auditability. Moreover, practical considerations such as multilingual support (Sinhala, Tamil, English), temporal legal variations, and on-premises deployment are often overlooked.

Overall, while existing systems provide strong foundational capabilities, they fall short in delivering a cost-effective, explainable, and domain-specific solution suitable for Sri Lankan deed verification.

1.2.3 Identified Research Gaps

Based on the analysis of existing literature and current technological solutions, several critical research gaps have been identified in the application of AI for legal document processing, particularly in the context of Sri Lankan deed verification.

Firstly, there is a lack of domain-specific datasets for Sri Lankan legal documents. Existing studies rely on datasets from contracts or case law in foreign jurisdictions, which do not reflect the structure, terminology, or legal requirements of local deeds. The absence of annotated datasets limits the development and benchmarking of accurate models.

Secondly, local legal rules are not formalized into machine-readable templates. Deeds in Sri Lanka require specific clauses such as attestation, chain of title, acceptance conditions, and mortgage release terms. However, these are not systematically encoded into AI systems, reducing their ability to perform reliable validation.

Another major gap is the inability to handle real-world document conditions. Most research assumes clean, digital inputs, whereas actual deeds often involve low-quality scans, skewed layouts, and handwritten elements. This reduces the effectiveness of existing models in practical scenarios.

Furthermore, there is a lack of explainability and evidence-linked outputs. Current systems typically provide predictions or risk scores without identifying the exact text spans that triggered those decisions. This limits trust and usability for legal professionals who require verifiable evidence.

The over-reliance on large language models (LLMs) also presents challenges. These models are expensive, require significant computational resources, and raise privacy concerns. The potential of small language models (SLMs) combined with rule-based approaches remains underexplored.

Additionally, legal grounding is insufficient, as outputs are rarely linked to relevant statutes or regulatory guidelines. There is also no standardized evaluation framework for Sri Lankan deeds, including issue taxonomies and severity classifications.

Other gaps include the lack of temporal awareness, where systems fail to consider laws applicable at the time of document execution, and the absence of human-in-the-loop mechanisms for incorporating expert feedback. Finally, end-to-end workflow integration, privacy-preserving deployment, and fairness across languages and document variations are not adequately addressed in existing research.

These gaps collectively highlight the need for a practical, explainable, and domain-specific solution tailored to Sri Lankan legal requirements.

1.2.4 Proposed Solution

This research proposes the development of an AI-based legal assistance system designed to provide accessible, affordable, and reliable legal guidance for users in Sri Lanka. The system focuses on Property Law and Family Law, which are commonly encountered yet complex for ordinary citizens. To achieve this, a structured dataset will be created by collecting legal information from books, documents, and expert sources, and converting it into a machine-readable format. SLM will then be fine-tuned using efficient techniques such as LoRA and Unsloth, enabling high performance with low computational cost.

To improve accuracy and reliability, the system will integrate an Agentic RAG architecture. This approach allows the model to retrieve relevant legal information from trusted sources, generate responses, and validate outputs before presenting them to users. The system is designed to provide step-by-step legal guidance in a clear and user-friendly manner, helping users understand procedures and make informed decisions. Additionally, it ensures that outputs are based on authorized data, reducing the risk of incorrect information. Finally, the solution will be deployed as a web-based application, offering a smooth user experience and enabling individuals to access legal assistance quickly and efficiently without depending entirely on professional lawyers.

1.3 Research Problem

The verification of legal deeds in Sri Lanka remains a critical yet challenging task due to the complexity, variability, and importance of such documents in property transfer and inheritance. Deeds such as Powers of Attorney, Sale Deeds, Gift Deeds, Testamentary Deeds, and Mortgage Deeds must adhere to strict legal requirements, including the presence of mandatory clauses, accurate identification of parties, clear property descriptions, and proper attestation. Any inconsistencies, omissions, or ambiguities in these documents can result in legal disputes, financial loss, and prolonged litigation.

Currently, the verification process is predominantly manual, carried out by legal professionals such as notaries and conveyancers. This process is time-consuming, labor-intensive, and prone to human error, especially when dealing with lengthy documents, repetitive checks, and poor-quality scanned copies. Moreover, the increasing volume of legal documentation further amplifies the need for efficient and reliable verification mechanisms.

Although recent advancements in Artificial Intelligence (AI), particularly in Natural Language Processing (NLP), offer potential solutions for automating legal document analysis, existing systems are not well-suited for this domain. Many rely on large language models (LLMs), which are computationally expensive, require cloud-based infrastructure, and raise concerns regarding data privacy and security. Additionally, these systems are typically trained on non-local datasets, making them less effective in capturing the specific legal structures and requirements of Sri Lankan deeds.

Furthermore, current approaches lack integration of domain-specific rules, evidence-based outputs, and end-to-end processing capabilities. They often fail to provide clear explanations linked to specific sections of the document, which is essential for legal validation and user trust. Challenges such as handling noisy scanned inputs, supporting multilingual content, and ensuring compliance with local legal standards remain largely unaddressed.

Therefore, the core research problem is to design and develop a cost-effective, explainable, and domain-specific AI system capable of automatically verifying Sri Lankan legal deeds. The system must accurately classify deed types, extract relevant information, enforce legal template rules, detect anomalies, and generate evidence-linked reports while operating efficiently in resource-constrained and privacy-sensitive environments.

1.3.1 Empirical Evidence

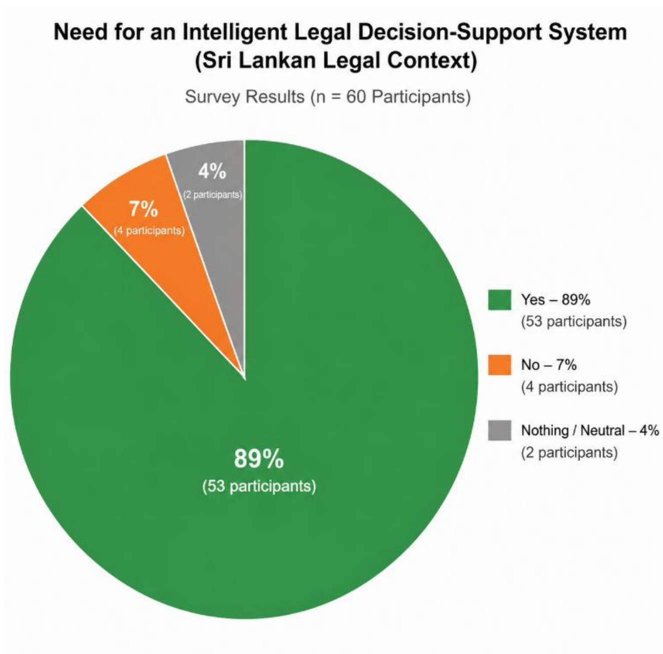


Figure 1:Survey of existing legal information

A survey was conducted with a total of 60 participants, including lawyers, law students, and members of the general public, to assess the need for an intelligent legal decision-support system tailored to the Sri Lankan context. The findings clearly indicate a strong demand for such a solution. Approximately 89% of respondents expressed a positive need for an AI-driven system to assist with legal processes, particularly in areas such as document verification, legal guidance, and decision support. This high percentage reflects widespread recognition of the challenges faced in traditional legal workflows.

In contrast, only 7% of participants indicated that such a system is not necessary, suggesting minimal resistance to technological adoption in the legal domain. Additionally, 4% of respondents remained neutral or uncertain, possibly due to limited awareness or familiarity with AI-based legal tools.

The survey results highlight key issues such as the time-consuming nature of manual legal verification, the risk of human error, and the lack of accessible legal support for non-experts. Overall, the findings strongly support the development of an intelligent, domain-specific legal system, confirming its relevance, demand, and potential impact within Sri Lanka.

1.3.2 Technological Limitations in Existing Systems

Despite rapid advancements in Artificial Intelligence (AI) and Natural Language Processing (NLP), existing legal document analysis systems exhibit several technological limitations that restrict their effectiveness in real-world deed verification, particularly within the Sri Lankan context.

One of the primary limitations is the dependence on large language models (LLMs). While LLMs offer strong capabilities in understanding complex legal text, they require high computational resources, significant memory, and cloud-based infrastructure. This makes them costly to deploy and unsuitable for low-resource environments. Additionally, reliance on cloud services raises data privacy and security concerns, especially when handling sensitive legal documents such as property deeds.

Another key limitation is the inability to effectively process long and unstructured documents. Legal deeds often span multiple pages with complex formatting, including tables, annotations, and irregular layouts. Although transformer-based models improve text understanding, they are constrained by token limits and require document segmentation, which can lead to loss of contextual continuity across sections.

Existing systems also struggle with poor-quality inputs, such as scanned documents with low resolution, skewed alignment, or faded text. Optical Character Recognition (OCR) errors significantly impact downstream tasks like information extraction and classification, reducing overall system reliability.

Furthermore, there is a lack of integration between rule-based reasoning and machine learning approaches. Most systems rely heavily on statistical or neural models without incorporating explicit legal rules, which are essential for validating structured documents like deeds. This results in systems that may understand text but fail to enforce legal correctness.

Another major limitation is the lack of explainability and evidence-based outputs. Many systems produce predictions or risk scores without linking them to exact text spans, making it difficult for legal professionals to verify results. This reduces trust and limits practical adoption.

Additionally, multilingual support remains inadequate, particularly for languages such as Sinhala and Tamil, which are commonly used in Sri Lankan legal documents. Most models are trained primarily on English data, leading to biased or inconsistent performance.

Finally, existing systems often lack end-to-end integration, focusing on isolated tasks such as OCR, classification, or entity recognition rather than providing a complete workflow. They also fail to address temporal legal variations, where rules may differ based on the date of document execution.

These technological limitations highlight the need for a lightweight, explainable, and domain-adapted system capable of addressing real-world constraints in deed verification.

1.4 Objectives

1.4.1 Main Objectives

The main objective of this research is to develop a cost-effective, explainable, and domain-specific AI-based system for the automated verification of Sri Lankan legal deeds.

Specifically, the system aims to:

Automatically identify and classify different types of deeds (e.g., Sale, Gift, Mortgage, Power of Attorney, Will)

Accurately extract key legal information such as parties, dates, property details, and financial terms

Apply rule-based template checks to ensure compliance with legal requirements

Use small language models (SLMs) to detect ambiguities and inconsistencies in legal text

Generate a concise summary and a detailed conformity report with page/line evidence

Provide a confidence score based on document quality and extraction accuracy

Overall, the objective is to reduce manual effort, improve accuracy, and enhance transparency in deed verification while ensuring suitability for low-resource and privacy-sensitive environments.

1.4.2 Specific Objectives

The proposed system is designed to achieve the following specific objectives:

1) To design and implement a document ingestion module capable of handling both digital PDFs and scanned deed documents with OCR support

2) To develop an information extraction system for identifying key legal entities such as names, dates, property details, and financial values

3) To build a deed classification model that accurately categorizes documents into predefined deed types

4) To create rule-based template frameworks for each deed type to validate mandatory clauses and structural requirements

5) To integrate small language models (SLMs) for detecting semantic issues such as ambiguity, inconsistency, and missing intent

6) To develop an agent-based routing system that assigns documents to specialized verification modules

7) To generate evidence-linked outputs, including page and line references for all detected issues

8) To compute a conformity score and confidence score for each processed document

9) To design an end-to-end pipeline that integrates all components from ingestion to final reporting

10) To ensure the system supports privacy-preserving, on-premises deployment suitable for real-world legal environments

2. METHODOLOGY

2.1 Key Technical Foundations of the Proposed System

This research adopts a hybrid, agent-based methodology that combines Small Language Models (SLMs), rule-based template matching, and document processing techniques to build a complete end-to-end deed verification system. The methodology is designed to reflect real-world legal workflows, ensuring that the system is both practical and scalable for Sri Lankan legal environments.

The overall approach follows a structured pipeline: document ingestion → validation → information extraction → deed classification → agent-based verification → scoring → report generation. Each stage is modular, allowing independent improvement and easy integration. The system is designed to handle both clean digital documents and noisy scanned inputs, ensuring robustness in real-world scenarios.

A key aspect of the methodology is the integration of deterministic rule-based checks with probabilistic SLM reasoning. Rule-based components enforce strict legal requirements, while SLMs provide contextual understanding to detect subtle issues such as ambiguity or inconsistent language. Additionally, the system emphasizes explainability, where every output is supported by evidence (page/line references), and efficiency, enabling deployment in low-resource, on-premises environments.

2.1.1 Small Language Models

Small Language Models (SLMs) are utilized in this research as a practical alternative to large language models, primarily due to their lower computational requirements, faster inference speed, and suitability for privacy-sensitive environments. Unlike large-scale models that require cloud-based infrastructure, SLMs can be deployed on local systems, making them ideal for legal applications involving confidential data such as property records. These models are fine-tuned using domain-specific legal texts to perform targeted tasks including semantic validation, ambiguity detection, and consistency checking within deed documents. By focusing on a narrow domain, SLMs can achieve high reliability without requiring extensive training data or resources. Additionally, the system employs carefully designed, task-specific prompts rather than general-purpose queries, which significantly reduces the risk of hallucinations and irrelevant outputs. This controlled usage ensures that SLMs provide precise, context-aware insights, making them an effective component in the hybrid deed verification framework.

2.1.2 Rule-Based Template Matching

Rule-based template matching forms the backbone of the proposed system by ensuring strict adherence to legal requirements in deed documents. Each type of deed—such as sale, gift, mortgage, or power of attorney—is associated with a predefined template that outlines mandatory clauses and structural elements based on Sri Lankan legal practices. These templates are translated into machine-readable rules, enabling automated validation of document content. For instance, the system checks for the presence of attestation details, verifies the chain of title, ensures acceptance clauses in gift deeds, and validates loan and security terms in mortgage documents. This deterministic approach guarantees that critical legal requirements are not overlooked, unlike purely statistical models that may miss such details. By combining these explicit rules with automated processing, the system achieves high accuracy and consistency in verification. This method also enhances explainability, as each validation step can be clearly traced back to a specific rule, making it highly suitable for legal applications.

2.1.3 Optical Character Recognition (OCR) and Document Processing

Optical Character Recognition (OCR) plays a crucial role in enabling the system to process real-world deed documents, which are often available as scanned images or low-quality digital copies. The OCR module is designed to extract textual content from various challenging inputs, including low-resolution images, skewed or rotated pages, faded ink, and multi-page PDF files. Given the variability in document quality, the system incorporates preprocessing techniques such as noise reduction, image alignment, and contrast enhancement to improve text extraction accuracy. Additionally, a quality scoring mechanism is implemented to assess the reliability of the OCR output. This score reflects factors such as text clarity, completeness, and recognition confidence, and is later used to influence the overall confidence level of the system's predictions. By integrating OCR with quality assessment, the system ensures that downstream processes—such as information extraction and validation—are informed by the reliability of the input data, thereby improving overall robustness.

2.1.4 Information Extraction Techniques

The system employs lightweight Natural Language Processing (NLP) techniques combined with pattern-based methods to extract structured information from unstructured deed text. Key legal entities such as names of parties, dates, property descriptions, financial amounts, and witness details are identified and organized into a structured format. This is achieved through a combination of regular expressions, keyword-based matching, and trained extraction models tailored to legal language patterns. The use of lightweight approaches ensures efficiency and reduces dependency on large datasets, making the system suitable for low-resource environments. Extracted information is then used for cross-validation across different sections of the document, helping to identify inconsistencies such as mismatched names or incorrect dates. This structured representation of data is essential for enabling rule-based validation and semantic analysis in later stages of the pipeline. Overall, the information extraction component serves as a critical bridge between raw document text and intelligent verification processes.

2.1.5 Deed Classification and Agent-Based Routing

The system incorporates a deed classification module that automatically identifies the type of legal document being processed. Using lightweight machine learning models and textual features, the system classifies documents into predefined categories such as Sale Deed, Gift Deed, Mortgage Deed, Power of Attorney, and Testamentary Deed. Once classified, the document is routed to a specialized verification agent designed specifically for that deed type. This agent-based architecture allows each module to focus on domain-specific validation rules and checks, thereby improving accuracy and efficiency. For example, a Gift Deed agent will prioritize acceptance clauses, while a Mortgage Deed agent will focus on loan and security terms. This targeted approach reduces system complexity by avoiding a one-size-fits-all model and enables better handling of document-specific requirements. Additionally, agent-based routing supports modular system design, allowing new deed types or rules to be added easily without affecting the overall system, ensuring scalability and maintainability.

2.2 Integrated Research Approach and System Methodology

The proposed system adopts an integrated, end-to-end research approach that combines document processing, machine learning, rule-based validation, and agent-based orchestration to address the challenges of Sri Lankan deed verification. Unlike traditional systems that focus on isolated tasks, this methodology ensures a complete workflow, aligning closely with real-world legal verification processes.

The system begins with document ingestion and validation, where input files in PDF or scanned formats are accepted. If the document is a scanned image, Optical Character Recognition (OCR) is applied to extract text, followed by a quality assessment to determine the reliability of the extracted content. This initial stage ensures that downstream processes are aware of input limitations.

Next, the system performs information extraction, where key legal entities such as names, dates, property descriptions, and financial details are identified using lightweight NLP techniques and pattern-based methods. The extracted data is structured and prepared for validation.

The deed classification module then identifies the document type and routes it to a specialized verification agent. Each agent is tailored to a specific deed category and applies both rule-based template checks and SLM-driven semantic analysis. Rule-based checks ensure compliance with mandatory legal requirements, while SLMs detect ambiguity, inconsistency, and contextual issues.

Following verification, the system performs cross-checking and scoring, assigning severity levels to detected issues and calculating an overall conformity score. A confidence score is also generated based on OCR quality and extraction completeness.

Finally, the system produces a comprehensive output, including a concise document summary and a detailed conformity report with page and line references. This integrated methodology ensures accuracy, explainability, and practical usability in real-world legal environments.

2.2.1 Agile Principles Applied in the Project

The development of the proposed deed verification system follows Agile principles to ensure flexibility, iterative improvement, and alignment with real-world legal requirements. Given the complexity of legal document processing and the variability of input data, an Agile approach enables continuous refinement of system components based on feedback and testing.

The project is structured into incremental development cycles (sprints), where each module—such as OCR processing, information extraction, classification, and verification—is developed, tested, and improved independently. This modular approach allows early identification of issues and ensures that each component functions correctly before integration into the full pipeline.

A key Agile principle applied is continuous feedback and improvement. Feedback from domain experts, such as legal practitioners or notaries, is incorporated to refine rule-based templates and improve system accuracy. This also supports the integration of human-in-the-loop mechanisms, enabling the system to learn from corrections and enhance performance over time.

The project also emphasizes adaptive design, allowing the system to evolve as new requirements emerge, such as adding support for new deed types or legal rules. Regular testing with real-world data ensures that the system remains robust and practical.

Overall, the use of Agile principles ensures that the system is flexible, scalable, and closely aligned with user needs, leading to a more reliable and effective solution.

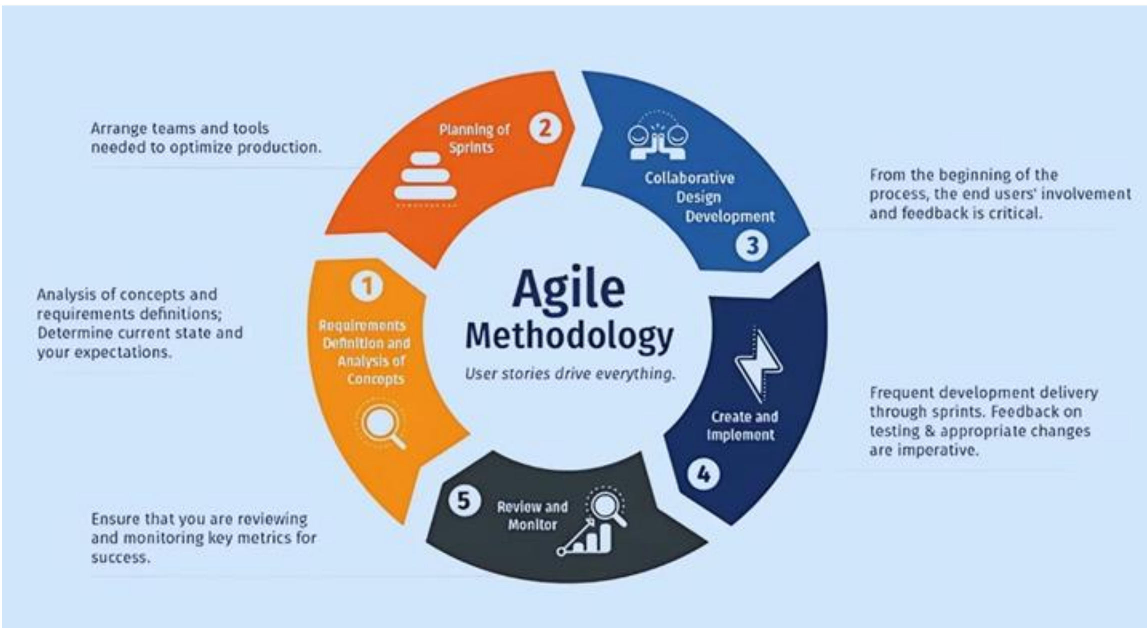


Figure 2: Agile Methodology

2.2.2 Feasibility Study and Planning

A comprehensive feasibility study was conducted to evaluate the practicality of developing and deploying the proposed deed verification system within the Sri Lankan context. The study considered technical, economic, operational, and legal feasibility to ensure that the solution is realistic and implementable in real-world environments.

From a technical perspective, the use of Small Language Models (SLMs) and lightweight NLP techniques was identified as feasible due to their low computational requirements and compatibility with on-premises systems. The availability of OCR tools and open-source libraries further supports the development of a complete document processing pipeline without the need for high-end infrastructure.

In terms of economic feasibility, the system is designed to minimize costs by avoiding expensive cloud-based large language models. This makes it suitable for small law firms, notaries, and organizations with limited resources. The use of open-source technologies also reduces development and deployment expenses.

The operational feasibility focuses on usability and integration into existing legal workflows. The system is designed to assist, rather than replace, legal professionals by providing evidence-linked outputs and easy-to-understand reports, ensuring smooth adoption.

From a legal and ethical perspective, the system prioritizes data privacy by supporting local deployment and secure handling of sensitive documents.

Planning was carried out in phases, including requirement analysis, system design, module development, integration, and testing, ensuring a structured and manageable development process.

2.2.3 Requirement Gathering and Analysis

The requirement gathering and analysis phase was conducted systematically to ensure that the proposed deed verification system aligns with real-world legal practices and user needs in Sri Lanka. This phase involved identifying both functional and non-functional requirements through a combination of domain study, literature review, and analysis of typical deed documents.

From a functional perspective, the system must be capable of handling end-to-end processing of legal deeds. Key requirements include document ingestion (PDF and scanned formats), OCR-based text extraction, identification and classification of deed types, extraction of critical legal fields, application of rule-based template checks, and SLM-based semantic analysis. Additionally, the system must generate a concise summary, a conformity report with severity levels, and evidence-linked outputs (page and line references), along with a confidence score indicating result reliability.

From a non-functional perspective, requirements focus on performance, usability, security, and scalability. The system must operate efficiently on low-resource machines, ensuring fast processing without reliance on high-end infrastructure. Data privacy and security are critical, requiring support for on-premises deployment and secure handling of sensitive legal documents. The system should also be user-friendly, providing clear and interpretable outputs for legal professionals.

Furthermore, special attention was given to real-world constraints, such as handling noisy scanned documents, supporting multilingual content (Sinhala, Tamil, and English), and adapting to variations in deed formats. This comprehensive requirement analysis forms the foundation for designing a practical, reliable, and domain-specific solution.

2.2.4 Research Design and Methodological Framing

The research design of this study follows a design science and system development approach, focusing on building and evaluating a practical solution to a real-world problem in Sri Lankan deed verification. The methodology is framed to combine both engineering design principles and applied research, ensuring that the system is not only theoretically sound but also practically implementable.

The study adopts a modular and iterative design, where the system is divided into key components such as document ingestion, OCR processing, information extraction, classification, rule-based validation, SLM-based reasoning, and report generation. Each module is developed and tested independently before being integrated into a complete end-to-end pipeline. This structured approach improves reliability and allows easier debugging and enhancement.

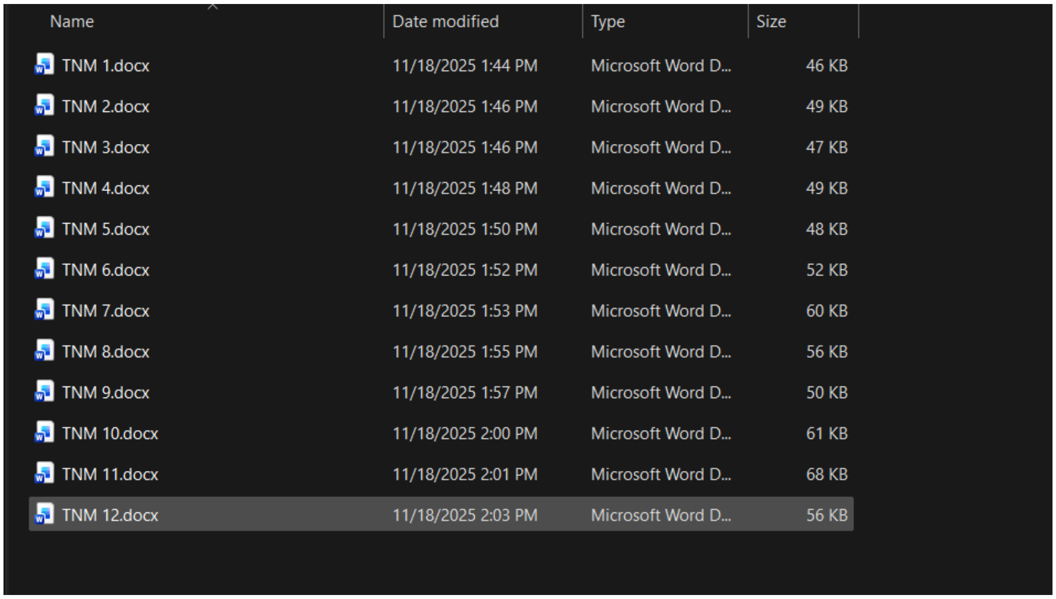
A hybrid methodological framework is used, combining rule-based techniques (for deterministic legal validation) with machine learning methods (for semantic understanding and pattern recognition). This ensures both accuracy and flexibility in handling structured and unstructured aspects of legal documents.

The research also incorporates experimental evaluation, where the system is tested on both clean digital documents and noisy scanned inputs to assess performance under realistic conditions. Key evaluation aspects include accuracy, explainability, processing efficiency, and robustness.

Overall, this research design ensures a balanced integration of theory, system development, and practical validation, making the proposed solution both effective and applicable in real-world legal environments.

2.2.5 Data Collection and Source Preparation

The data collection process for this research was primarily based on authoritative Sri Lankan legal resources, including law books and professional guides used in drafting and reviewing deeds. These sources provide detailed explanations of legal structures, mandatory clauses, and standard practices followed by notaries and conveyancers. By analyzing these materials, relevant textual patterns, legal terminology, and document structures were identified and compiled. Since publicly available datasets for Sri Lankan deeds are not readily accessible, this approach ensured that the dataset reflects realistic and domain-specific legal knowledge. The collected content includes examples of various deed types such as Sale Deeds, Gift Deeds, Powers of Attorney, Mortgage Deeds, and Testamentary Deeds. Special attention was given to capturing variations in language, clause structures, and formatting styles commonly observed in practice. This manual and knowledge-driven data collection strategy ensures that the dataset is both representative and aligned with real-world legal requirements, forming a strong foundation for model training and system development.



Name	Date modified	Type	Size
TNM 1.docx	11/18/2025 1:44 PM	Microsoft Word D...	46 KB
TNM 2.docx	11/18/2025 1:46 PM	Microsoft Word D...	49 KB
TNM 3.docx	11/18/2025 1:46 PM	Microsoft Word D...	47 KB
TNM 4.docx	11/18/2025 1:48 PM	Microsoft Word D...	49 KB
TNM 5.docx	11/18/2025 1:50 PM	Microsoft Word D...	48 KB
TNM 6.docx	11/18/2025 1:52 PM	Microsoft Word D...	52 KB
TNM 7.docx	11/18/2025 1:53 PM	Microsoft Word D...	60 KB
TNM 8.docx	11/18/2025 1:55 PM	Microsoft Word D...	56 KB
TNM 9.docx	11/18/2025 1:57 PM	Microsoft Word D...	50 KB
TNM 10.docx	11/18/2025 2:00 PM	Microsoft Word D...	61 KB
TNM 11.docx	11/18/2025 2:01 PM	Microsoft Word D...	68 KB
TNM 12.docx	11/18/2025 2:03 PM	Microsoft Word D...	56 KB

Figure 3:All the data collected data

2.2.6 Dataset Creation for Deed Classification

Using the collected legal materials, a structured dataset consisting of approximately 1,500 samples was created for the purpose of deed type classification. Each sample represents a segment or full text of a deed categorized into predefined classes such as Sale, Gift, Mortgage, Power of Attorney, and Testamentary documents. The dataset was carefully labeled based on domain knowledge to ensure accuracy and consistency. Preprocessing steps were applied to clean and standardize the text, including removing noise, normalizing formatting, and segmenting long documents where necessary. This dataset was then used to train a classification model capable of automatically identifying the type of deed based on textual features. The use of a moderate-sized, high-quality dataset ensures efficient training while maintaining relevance to the domain. This classification component plays a crucial role in the system by enabling accurate routing of documents to specialized verification agents, thereby improving overall system performance and reliability.

2.2.7 Prompt Design and AI Agent Preparation

In addition to dataset creation, the collected legal sources were used to design prompts and develop the processing logic for the AI-based verification agent. These prompts are carefully crafted to guide small language models (SLMs) in identifying potential issues within deed documents. The prompts focus on specific legal aspects such as clause completeness, ambiguity detection, consistency of information, and adherence to legal requirements. By grounding the prompts in domain-specific knowledge derived from legal books, the system ensures more accurate and context-aware responses. Furthermore, a processing pipeline was developed within the agent framework to integrate prompt execution with rule-based validation and information extraction. This allows the system to perform structured analysis while also leveraging semantic understanding. The combination of prompt engineering and rule-based logic enhances the system's ability to detect both explicit and implicit issues, making the verification process more robust, explainable, and aligned with real-world legal practices..

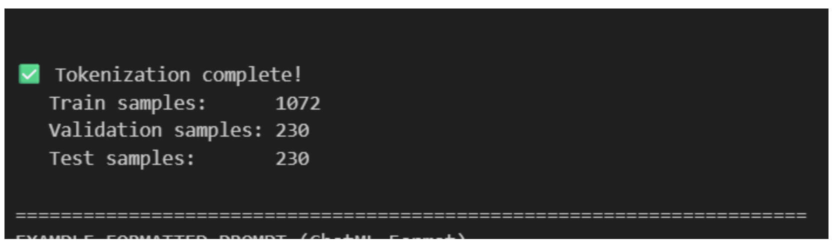


Figure 4:Splitting, and Schema Validation

2.2.8 Model Fine-Tuning Strategy

The model fine-tuning strategy in this research focuses on adapting a small language model (SLM) to accurately perform deed classification within the Sri Lankan legal domain. Instead of training a model from scratch, a pre-trained lightweight transformer-based model is fine-tuned using domain-specific data to improve efficiency and performance.

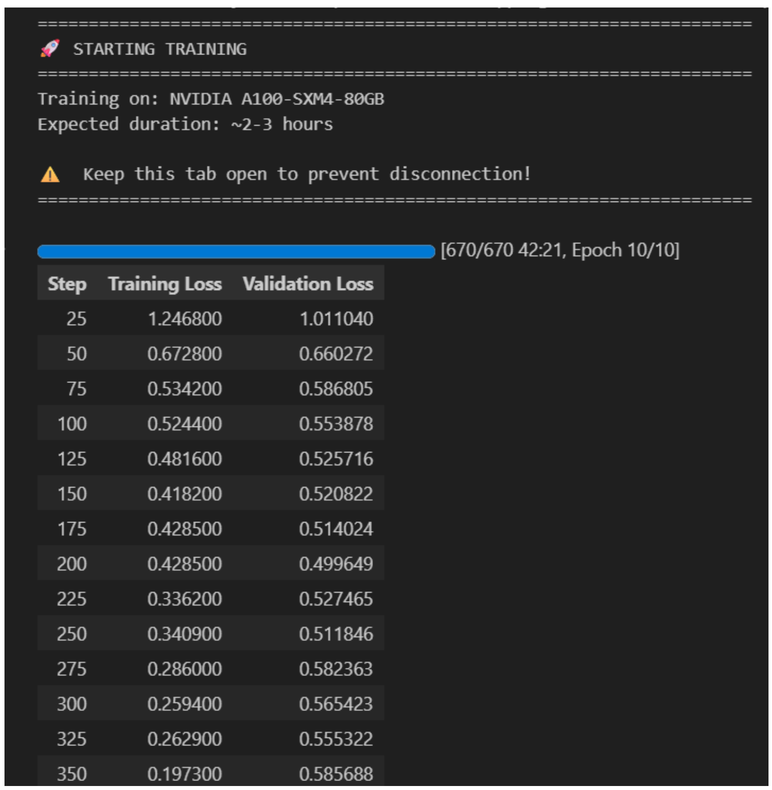
A dataset consisting of 1,500+ labeled deed samples was used for training. Each sample includes structured input-output pairs, where the input represents the deed content and the output corresponds to the deed type. The dataset was divided into training, validation, and testing subsets to ensure proper evaluation and generalization.

To optimize resource usage, Low-Rank Adaptation (LoRA) was employed. This technique allows the model to learn task-specific patterns by updating only a small subset of parameters, significantly reducing computational cost while maintaining high performance.

The model was fine-tuned using instruction-based prompt formatting, where each input follows a structured pattern to guide the model toward accurate classification.

Training was conducted using controlled hyperparameters, including a moderate learning rate, batch size optimization, gradient accumulation, and regularization techniques such as dropout and weight decay to prevent overfitting. Evaluation metrics such as accuracy and F1-score were used to measure performance.

The fine-tuned model achieved high classification accuracy and was integrated into the system for deed-type identification and agent routing, forming a critical component of the overall pipeline.



```
======================================================================
🚀 STARTING TRAINING
======================================================================
Training on: NVIDIA A100-SXM4-80GB
Expected duration: ~2-3 hours

⚠️  Keep this tab open to prevent disconnection!
======================================================================
```

[670/670 42:21, Epoch 10/10]

Step	Training Loss	Validation Loss
25	1.246800	1.011040
50	0.672800	0.660272
75	0.534200	0.586805
100	0.524400	0.553878
125	0.481600	0.525716
150	0.418200	0.520822
175	0.428500	0.514024
200	0.428500	0.499649
225	0.336200	0.527465
250	0.340900	0.511846
275	0.286000	0.582363
300	0.259400	0.565423
325	0.262900	0.555322
350	0.197300	0.585688

Figure 5:model training

2.2.10 Agent Development Strategy

The agent development strategy in this research is designed to build a modular, scalable, and domain-specific multi-agent system for legal deed verification. The approach follows a layered and inheritance-based design, where each agent is developed with a clear responsibility and can operate independently while contributing to the overall workflow.

At the core of the strategy is the use of a BaseValidator class, which defines common functionalities such as input handling, rule execution, output formatting, and error management. All deed-specific agents inherit from this base class, ensuring consistency in structure while allowing customization of validation logic. This object-oriented approach simplifies development, maintenance, and future expansion.

Each validator agent is developed using a combination of rule-based logic and prompt-driven reasoning. Domain knowledge from Sri Lankan legal sources is encoded into both explicit validation rules and structured prompts, enabling agents to detect missing clauses, inconsistencies, and ambiguous language. Prompts are carefully designed to produce structured outputs, improving interpretability and downstream processing.

The strategy also emphasizes incremental and iterative development, where agents are built and tested individually before integration into the orchestration layer. This allows early detection of errors and continuous refinement based on test results and expert feedback.

To ensure robustness, agents are equipped with fallback mechanisms and error handling, enabling graceful degradation in cases of low confidence or unexpected input. Additionally, structured schemas are used to standardize outputs across all agents.

Overall, this strategy enables the creation of a flexible, explainable, and extensible agent-based system tailored for real-world legal applications.

2.2.11 System Integration, Architecture, and Observability

The proposed system is designed as an integrated, end-to-end architecture that combines multiple components into a unified workflow for legal deed verification. The architecture follows a modular and layered design, ensuring that each component—such as OCR processing, classification, validation agents, and reporting—can function independently while remaining seamlessly connected within the overall pipeline.

System integration is achieved through a central orchestration mechanism, which manages the flow of data between components. The workflow begins with document ingestion and text extraction, followed by classification, agent-based validation, and finally output generation. Each stage communicates through structured data formats, enabling smooth data transfer and reducing dependency issues. This design allows easy updates or replacement of individual modules without affecting the entire system.

The architecture is built to support scalability and flexibility, allowing additional deed types, validation rules, or models to be incorporated with minimal changes. It also supports on-premises deployment, ensuring data privacy and compliance with legal requirements.

Observability is a key aspect of the system, ensuring transparency, monitoring, and debugging capabilities. The system maintains detailed logs, step-by-step execution traces, and metadata records for each processed document. These logs capture classification decisions, validation results, detected issues, and confidence scores. Such traceability enables users to audit system behavior and understand how conclusions are derived.

Additionally, error handling and monitoring mechanisms are implemented to detect failures or inconsistencies during processing. This ensures reliability and allows developers to quickly identify and resolve issues.

Overall, the integrated architecture and observability features ensure that the system is robust, transparent, and maintainable, making it suitable for real-world legal applications.

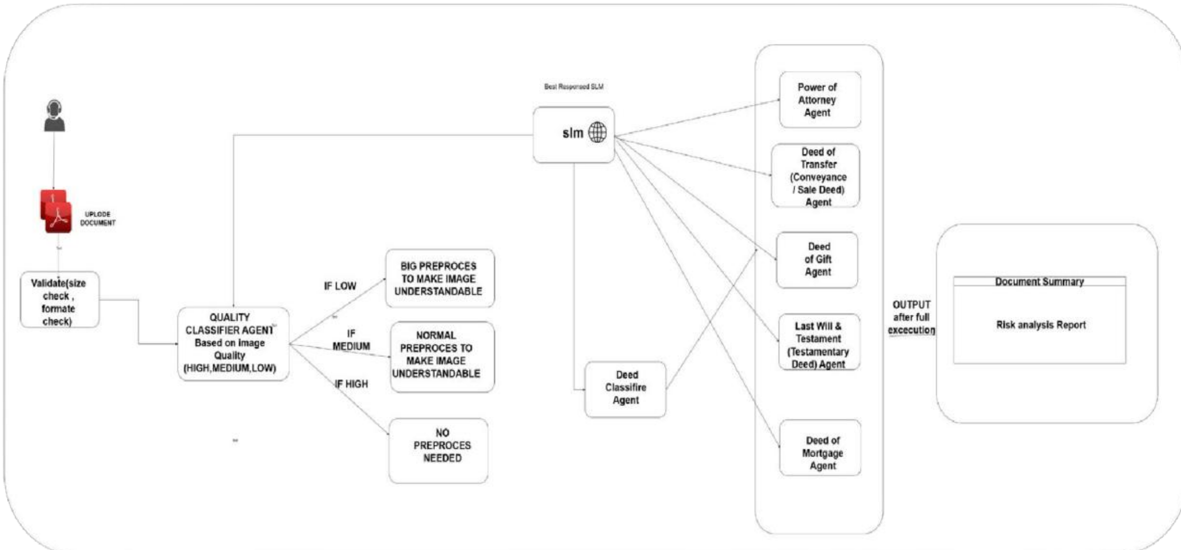


Figure 6:System Architecture Diagram

2.2.12 Evaluation, Reliability Controls, and Iterative Refinement

The proposed system incorporates a comprehensive evaluation and reliability framework to ensure accuracy, robustness, and trustworthiness in legal deed verification. Evaluation is conducted at multiple levels, including classification performance, information extraction quality, and validation accuracy. Standard metrics such as accuracy, precision, recall, and F1-score are used to assess the performance of the deed classification model, while validation outputs are evaluated based on correctness, completeness, and alignment with legal rules.

To enhance reliability, the system integrates several control mechanisms. These include confidence scoring based on OCR quality, extraction completeness, and model prediction certainty. When confidence levels fall below predefined thresholds, fallback strategies are triggered to maintain output reliability. Additionally, rule-based validation acts as a deterministic safeguard, ensuring that critical legal requirements are always enforced regardless of model behavior.

The system also emphasizes explainability, where all detected issues are linked to specific page and line references, enabling easy verification by users. Structured output schemas further ensure consistency and reduce ambiguity in results.

An iterative refinement approach is adopted to continuously improve system performance. Feedback from testing and domain experts is used to refine rules, prompts, and model behavior. Errors identified during evaluation are analyzed and used to update validation logic and training data.

Overall, this combination of evaluation metrics, reliability controls, and continuous improvement ensures that the system remains accurate, stable, and adaptable to real-world legal scenarios.

2.2.13 Project Timeline and Gantt Chart

The project was planned and executed using a structured timeline to ensure systematic development, testing, and refinement of the proposed deed verification system. The entire research was divided into multiple phases, each focusing on a specific component of the system. This phased approach helped in maintaining clarity, tracking progress, and ensuring timely completion of tasks.

The project began with requirement analysis and literature review, where existing research gaps and system requirements were identified. This was followed by the data collection and dataset preparation phase, where legal documents and law books were used to create training data for the classification model.

The next phase involved model development and fine-tuning, including training the SLM classifier using LoRA techniques. Parallel to this, the system design and agent development phase was carried out, where the multi-agent architecture and validation logic were implemented.

Subsequently, the project moved to system integration, combining all modules such as OCR, classification, validation agents, and reporting into a unified pipeline. This was followed by testing and evaluation, where the system was validated using both clean and noisy documents. Finally, documentation and refinement were completed, incorporating feedback and improving system performance.

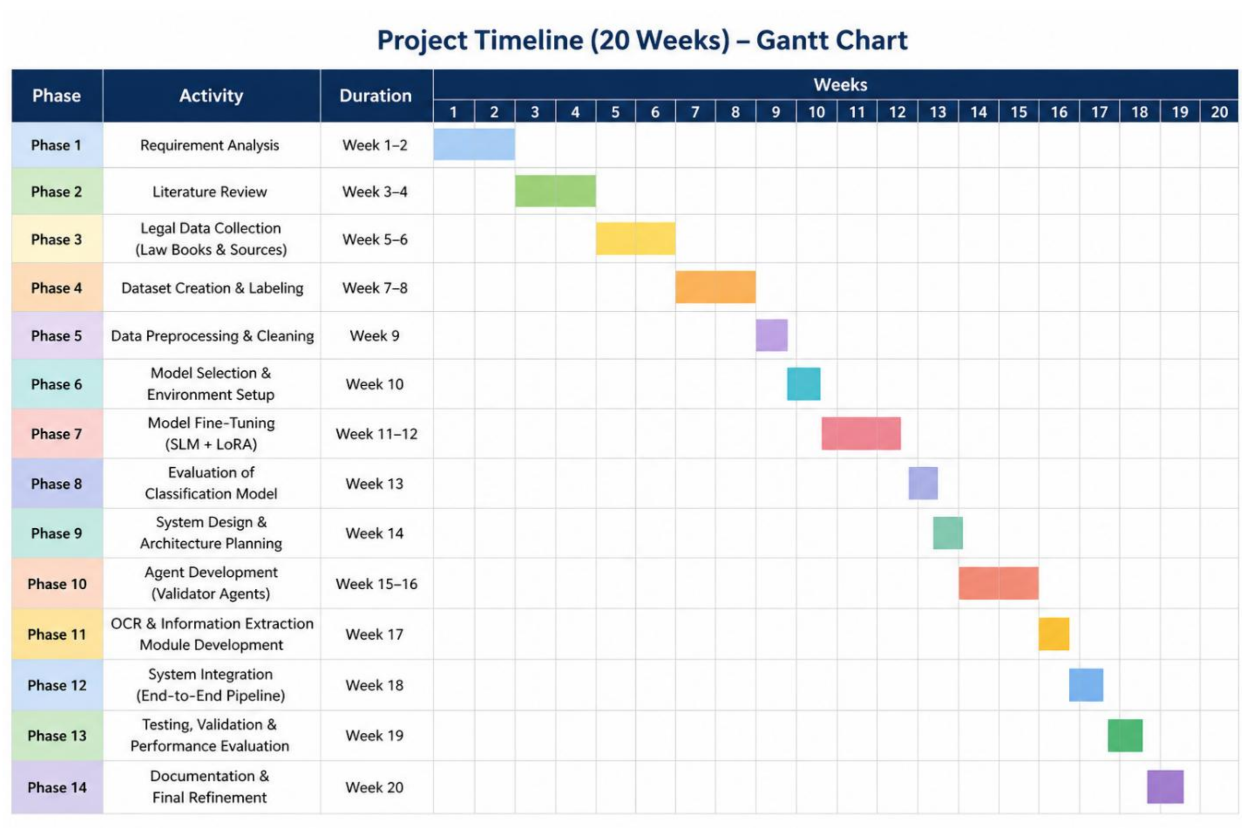


Figure 7:Gantt Chart

2.3 Summary of Methodology

This research follows a structured, hybrid methodology to develop an AI-based deed verification system for the Sri Lankan legal context. The proposed approach combines small language models (SLMs), rule-based template matching, OCR-based document processing, and multi-agent system design to create a practical, explainable, and cost-effective solution.

The methodology begins with data collection and preparation, where Sri Lankan legal books and deed drafting materials are used to create a dataset for deed classification. Approximately 1,500 deed samples are prepared and labelled according to key deed types, including Power of Attorney, Sale Deed, Gift Deed, Testamentary Deed, and Mortgage Deed. These data are cleaned, structured, and used to fine-tune a lightweight classification model.

A fine-tuned Qwen-based small language model with LoRA adaptation is used as the deed classification component. This model identifies the deed type and routes the document to the relevant validation agent. The use of SLMs reduces computational cost and supports privacy-preserving deployment compared with large cloud-based models.

The system follows an end-to-end processing pipeline. First, documents are ingested as PDFs or scanned files. If needed, OCR is applied to extract text, and a quality score is generated to measure extraction reliability. Next, key information such as parties, dates, property descriptions, financial values, witnesses, and notary details is extracted using lightweight NLP and pattern-based techniques.

After classification, the document is routed to a deed-specific validator agent. Each agent applies machine-checkable legal template rules relevant to its deed type. These rules check important legal requirements such as attestation details, chain of title, acceptance language in gift deeds, executor details in testamentary deeds, and loan/security terms in mortgage deeds. In addition, focused SLM prompts are used to detect semantic issues such as ambiguity, unclear wording, name mismatches, missing intent, and inconsistent clauses.

The system then produces two major outputs: a concise document summary and a conformity report. The conformity report highlights detected issues with severity levels, conformity score, confidence score, and exact page/line evidence. Confidence is calculated using OCR quality, extraction completeness, and model certainty.

Overall, the methodology provides a complete workflow from document intake to final report generation. It is designed to be modular, explainable, low-cost, and suitable for real-world legal practice in Sri Lanka.

2.4. Commercialization aspects of the product

2.4.1 Product Overview and Value Proposition

The proposed system is a legal-tech product designed for automated deed verification tailored specifically to the Sri Lankan legal environment. It integrates small language models (SLMs), rule-based validation, and multi-agent architecture to provide an efficient, explainable, and cost-effective solution for analyzing legal deeds.

The core functionality of the product includes automatic deed classification, key information extraction, legal validation, and generation of structured outputs, such as concise summaries and conformity reports with evidence-linked findings. Unlike existing solutions that rely heavily on large, expensive cloud-based models, this system is designed for on-premises deployment, ensuring data privacy and compliance with legal requirements.

The primary value proposition of the product lies in its ability to

Reduce manual effort and verification time for legal professionals

Improve accuracy and consistency in identifying legal issues

Provide transparent, evidence-based outputs with page and line references

Operate in low-resource environments with minimal computational cost

Ensure compliance with Sri Lankan legal standards through rule-based validation

By combining explainability, affordability, and domain specificity, the system offers a practical alternative to traditional manual verification and existing AI solutions.

2.4.2 Target Market and Users

The target market for the proposed product includes professionals and organizations involved in legal documentation, property transactions, and compliance verification within Sri Lanka.

The primary users are:

Notaries and conveyancers, who are responsible for drafting and verifying deeds

Law firms and legal practitioners, who require efficient tools for document analysis and validation

Real estate companies and property developers, who frequently handle property transfer documents

Financial institutions, such as banks and mortgage providers, that need to verify legal documents related to loans and securities

Government and land registry offices, where large volumes of deed documents are processed

Additionally, the system can be extended to support individual users (citizens) seeking basic legal guidance or document verification assistance.

The product is particularly valuable in environments where high document volume, limited resources, and strict privacy requirements exist. Its ability to run on local systems makes it suitable for organizations that cannot rely on cloud-based solutions due to data sensitivity.

Overall, the system targets a niche but critical market segment in Sri Lanka, offering a scalable solution that addresses real-world legal challenges.

2.4.3 Deployment Model and Architecture for Commercial Use

The proposed system is designed with a flexible and privacy-aware deployment model to support real-world commercial use in the Sri Lankan legal domain. Given the sensitivity of legal documents such as property deeds, the system primarily adopts an on-premises deployment architecture, ensuring that all data processing occurs within the organization's secure environment. This approach minimizes data exposure, complies with legal and regulatory requirements, and builds trust among users such as law firms, notaries, and financial institutions.

The architecture follows a modular microservice-based design, where key components—such as OCR processing, classification model, validation agents, and reporting services—are deployed as independent but interconnected services. This allows organizations to scale individual components based on demand and update modules without disrupting the entire system.

For enhanced flexibility, the system also supports a hybrid deployment model, where non-sensitive operations (e.g., model updates or external validation fallback) can be optionally integrated with cloud-based services. This ensures a balance between performance, scalability, and cost efficiency.

The system is typically exposed through a web-based interface or API layer, enabling seamless integration with existing legal workflows, document management systems, or enterprise applications. Secure access controls, authentication mechanisms, and audit logging are implemented to ensure data integrity and traceability.

Overall, the deployment architecture is designed to be secure, scalable, and adaptable, making it suitable for commercial adoption across various legal and enterprise environments.

2.4.4 Legal, Ethical, and Regulatory Considerations

The deployment and commercialization of the proposed deed verification system require careful consideration of legal, ethical, and regulatory factors, particularly due to the sensitive nature of property and legal documents in Sri Lanka.

From a legal perspective, the system must ensure compliance with Sri Lankan laws governing data protection, property registration, and legal documentation. Since deeds contain confidential personal and financial information, the system is designed to support on-premises deployment, ensuring that data remains within the organization's secure infrastructure. Additionally, the system does not replace legal professionals but acts as a decision-support tool, meaning that final legal responsibility remains with qualified practitioners such as notaries and lawyers.

Ethically, the system emphasizes transparency, fairness, and accountability. All outputs are designed to be explainable, with issues linked to specific page and line references, enabling users to verify results. The system also incorporates confidence scores to communicate uncertainty, preventing over-reliance on automated decisions. Care is taken to minimize bias, especially in handling multilingual documents (Sinhala, Tamil, and English) and varying document formats.

From a regulatory standpoint, the system must align with standards related to digital record handling, auditability, and compliance monitoring. Features such as secure storage, access control, and audit logs are implemented to ensure traceability and accountability of all operations.

Overall, these considerations ensure that the system is legally compliant, ethically responsible, and suitable for adoption in real-world legal environments.

2.4.5 Competitive Advantage

The proposed deed verification system offers several distinct competitive advantages over existing legal AI solutions, particularly within the Sri Lankan context. Unlike most current systems that are designed for contracts or case law in foreign jurisdictions, this solution is domain-specific, tailored explicitly for Sri Lankan deeds and local legal practices. This localization ensures higher relevance, accuracy, and practical usability.
One of the key advantages is the use of small language models (SLMs) instead of large language models. This significantly reduces computational cost, improves processing speed, and enables on-premises deployment, addressing critical concerns related to data privacy and security. As a result, the system is more accessible to small and medium-sized legal firms that cannot afford expensive cloud-based AI solutions.
Another major strength is the hybrid validation approach, which combines rule-based template checks with SLM-based semantic analysis. This ensures both deterministic accuracy (through

legal rules) and contextual understanding (through AI), reducing errors and improving reliability. Additionally, the system provides evidence-linked outputs, including page and line references, which enhances transparency and builds trust among legal professionals.

The multi-agent architecture further improves modularity and scalability, allowing new deed types or validation rules to be added without modifying the entire system. This makes the solution adaptable to evolving legal requirements.

Overall, the system's combination of cost efficiency, explainability, domain specificity, and privacy-focused design provides a strong competitive edge in the legal-tech market.

2.5 Project Requirements

2.5.1 Functional requirements

The functional requirements define the core capabilities that the proposed deed verification system must perform to meet its objectives. The system should support end-to-end processing of legal deed documents, starting from input to final output generation.

Firstly, the system must allow users to upload documents in multiple formats, including PDF and scanned images. It should automatically perform text extraction using OCR when required. The system must then classify the deed type into predefined categories such as Sale Deed, Gift Deed, Mortgage Deed, Power of Attorney, and Testamentary Deed.

The system should also extract key legal information, including names of parties, dates, property details, financial values, witnesses, and notary information. Following extraction, it must apply rule-based validation checks specific to each deed type to ensure compliance with legal requirements.

In addition, the system should utilize SLM-based analysis to detect semantic issues such as ambiguity, inconsistency, and missing clauses. It must generate two primary outputs: a concise document summary and a detailed conformity report highlighting detected issues with severity levels and page/line evidence.

Finally, the system should compute and display conformity scores and confidence scores, and support integration with external systems through APIs if required.

2.5.2 Non-functional requirements

The non-functional requirements define the quality attributes and constraints that ensure the system performs efficiently, securely, and reliably in real-world environments.

The system must ensure high performance and efficiency, with fast processing times even on low-resource machines. It should be optimized for on-premises deployment, minimizing dependency on cloud infrastructure to maintain data privacy.

Security and data protection are critical, as the system handles sensitive legal documents. It must implement secure data storage, user authentication, and access control mechanisms. Additionally, the system should maintain audit logs to track all processing steps for transparency and compliance.

The system must provide high reliability and robustness, handling noisy scanned documents, OCR errors, and unexpected inputs gracefully. It should include fallback mechanisms and error handling to ensure continuity of operation.

Usability is also important; the system should present outputs in a clear, structured, and user-friendly format that legal professionals can easily interpret. Furthermore, the system should support scalability and modularity, allowing new deed types, rules, or components to be added without major redesign.

Finally, the system should ensure fairness and multilingual support, accommodating documents in Sinhala, Tamil, and English to avoid bias and ensure consistent performance across different document types.

2.6 Testing and Implementation

2.6.1 Testing

Overview of Testing Phase

The testing phase of this research was designed to ensure the correctness, reliability, and robustness of the proposed deed verification system across all its components, including the fine-tuned model, API layer, workflow orchestration, and agent modules. A multi-level testing strategy was adopted, combining unit testing, integration testing, and system-level validation.

At the foundational level, testing focused on configuration consistency and data validation. Core components such as schema definitions, input constraints, and rule mappings were verified to ensure that the system correctly handles legal document inputs and enforces required formats. These tests confirmed that the system behaves predictably even before invoking model inference or external workflows .

The fine-tuned classification model was evaluated separately using standard machine learning metrics, including accuracy and F1-score, to ensure its effectiveness in identifying deed types. The model demonstrated high performance, making it reliable for routing documents to the correct validation agents.

At the application level, API testing was performed to validate endpoints, request handling, and response structures. Mocking techniques were used to isolate the workflow logic, ensuring that the system produces consistent outputs regardless of external dependencies. This allowed verification of error handling, input validation, and response formatting.

Additionally, workflow-level testing was conducted on the orchestration layer to validate conditional routing, error handling, and state transitions. This ensured that the system correctly manages different processing paths, such as handling incomplete data or routing failures.

Finally, agent-level testing verified the correctness of validator selection and prompt structures, ensuring that each deed type is processed by the appropriate agent. Overall, the testing phase confirms that the system is accurate, stable, and suitable for real-world deployment.

Fine-Tuned Model Performance Testing

The fine-tuned model performance testing was conducted to evaluate how accurately the small language model could classify Sri Lankan deed documents into the correct deed type. This testing stage is important because the classification model acts as the routing component of the overall deed verification system. If the model predicts the wrong deed type, the document may be sent to the wrong validator agent, which can affect the accuracy of the final conformity report.

The model was trained and tested using a dataset of more than 1,500 deed samples, created from Sri Lankan deed drafting materials. The dataset was divided into training, validation, and testing sets to measure generalization performance. The model was fine-tuned using LoRA adaptation, allowing efficient training without updating all parameters of the base model.

During testing, the model was evaluated using standard classification metrics such as accuracy, precision, recall, macro F1-score, and weighted F1-score. These metrics were selected because they provide a balanced understanding of model performance across different deed classes. Accuracy measures the overall correctness of predictions, while precision and recall show how well each class is identified. F1-score was especially important because it balances precision and recall.

The testing results showed strong classification performance, with the fine-tuned model achieving approximately 99.13% accuracy, 0.9900 macro F1-score, and 0.9913 weighted F1-score. These results indicate that the model can reliably distinguish between major deed types such as Power of Attorney, Deed of Transfer, Mortgage Deed, Deed of Gift, and Testamentary Deed.

Overall, the fine-tuned model demonstrated high reliability for deed-type classification and was suitable for integration into the agent-based validation pipeline. However, future testing should include more real-world scanned deeds, multilingual documents, and unseen deed formats to further validate robustness.

```
======================================================================
🎯 MODEL PERFORMANCE METRICS
======================================================================
Overall Accuracy:    99.13%
Macro F1-Score:      0.9900
Weighted F1-Score:   0.9913
======================================================================


======================================================================
CLASSIFICATION REPORT
======================================================================
                   precision    recall  f1-score   support

     Deed of Gift       0.94      1.00      0.97        32
 Deed of Transfer       1.00      1.00      1.00        43
             GIFT       1.00      0.94      0.97        35
         MORTGAGE       1.00      1.00      1.00        33
Power of Attorney       1.00      1.00      1.00        55
Testamentary Deed       1.00      1.00      1.00        32

         accuracy                           0.99       230
        macro avg       0.99      0.99      0.99       230
     weighted avg       0.99      0.99      0.99       230
```

Figure 8:Finetuned model evaluation

Core configuration and models Testing

The core configuration and models testing phase was conducted to ensure the correctness, consistency, and stability of the foundational components of the proposed deed verification system before executing full workflow operations. This stage focuses on validating configuration settings, data schemas, and model-related utilities without involving external APIs, orchestration layers, or runtime inference.

Initially, the system's configuration layer was tested to verify the integrity of critical parameters such as environment variables, API keys, and system settings. The configuration module ensures that all required values are properly loaded and accessible during runtime. Special attention was given to validating mappings such as deed type mappings, which link human-readable labels to internal system codes. These mappings are essential for correct routing and agent selection.

Next, data models and schema validation were tested using structured definitions (e.g., Pydantic models). These tests ensured that all inputs meet required constraints, such as

minimum text length, valid field types, and allowed values for enumerations like issue severity levels. This guarantees that only valid and well-formed data enters the processing pipeline.
Additionally, deterministic utility functions were tested to confirm consistent behavior. For example, scoring functions were validated to ensure correct and monotonic mapping from numeric values to grade levels, which is important for interpretability in legal outputs.
Finally, the prompt configuration and registry were verified to ensure that each supported deed type has a corresponding prompt template. This prevents runtime errors during agent execution.
Overall, this testing phase confirms that the system's core configuration, data structures, and model dependencies are reliable and internally consistent, forming a strong foundation for higher-level processing.

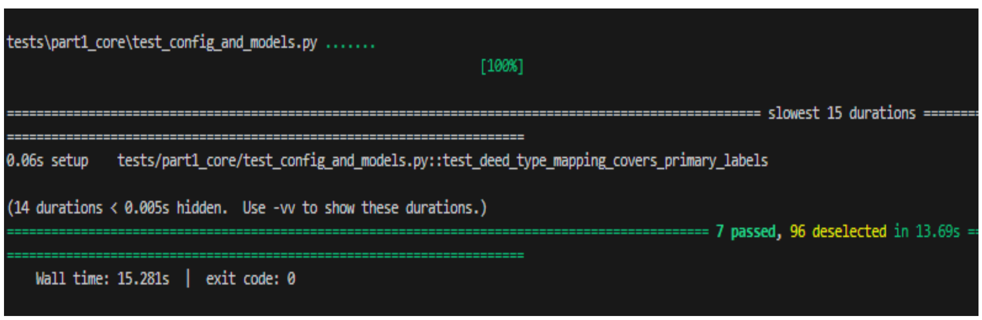


Figure 9:core configuration and model testing

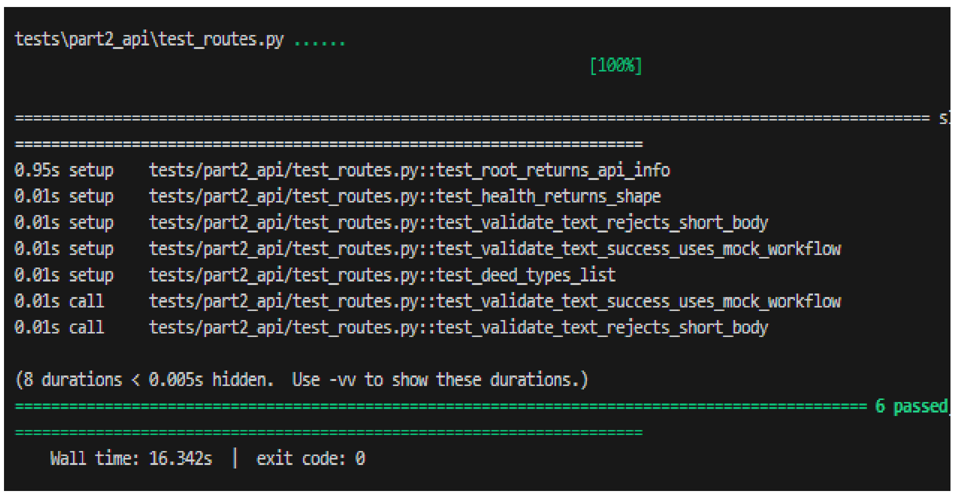
API Routes Testing

API routes testing was conducted to verify the correctness, stability, and reliability of the FastAPI application layer. This testing phase focused on checking whether the system endpoints correctly handle user requests, validate inputs, return proper responses, and manage errors.
The main endpoints tested included the root route, health check route, deed type listing route, and deed validation route. The root and health check endpoints confirmed that the application was running correctly and ready to receive requests. The deed type endpoint was tested to

ensure that all supported deed categories, such as Power of Attorney, Sale Deed, Gift Deed, Mortgage Deed, and Testamentary Deed, were returned correctly.

For the validation endpoint, test cases were created using mocked workflow responses. This allowed the API layer to be tested without depending on the full model, LangGraph workflow, or external API calls. The system was also tested with invalid inputs, such as deed text below the required minimum length, to confirm that proper error responses were returned.

Response formatting was also verified to ensure that successful and failed validation results followed a consistent JSON structure. Overall, the API routes testing confirmed that the application layer works correctly, handles errors properly, and provides stable communication between the user interface and the deed verification pipeline.



```
tests\part2_api\test_routes.py ......
                                                        [100%]

========================================================================================== s
==============================================================
0.95s setup    tests/part2_api/test_routes.py::test_root_returns_api_info
0.01s setup    tests/part2_api/test_routes.py::test_health_returns_shape
0.01s setup    tests/part2_api/test_routes.py::test_validate_text_rejects_short_body
0.01s setup    tests/part2_api/test_routes.py::test_validate_text_success_uses_mock_workflow
0.01s setup    tests/part2_api/test_routes.py::test_deed_types_list
0.01s call     tests/part2_api/test_routes.py::test_validate_text_success_uses_mock_workflow
0.01s call     tests/part2_api/test_routes.py::test_validate_text_rejects_short_body

(8 durations < 0.005s hidden.  Use -vv to show these durations.)
==================================================================================== 6 passed
==============================================================
    Wall time: 16.342s  |  exit code: 0
```

Figure 10:Api Rout testing

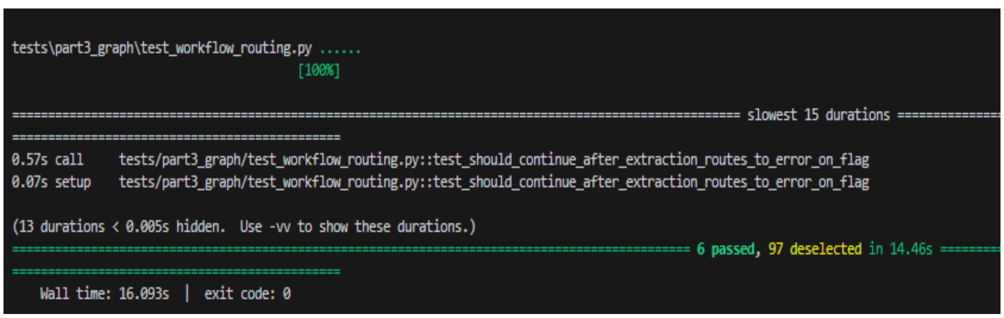
Agent Graph Workflow Routing Testing

Agent graph workflow routing testing was conducted to verify whether the LangGraph-based orchestration layer correctly manages the movement of documents through the deed verification pipeline. This testing phase focused on checking the conditional routing logic that controls how the system moves from one stage to another, such as text extraction, deed classification, validation, error handling,and output formatting.

The tests evaluated whether the workflow correctly handles both successful and failed processing conditions. For example, if text extraction fails, if the extracted deed text is too short, or if the classification result is missing, the system should stop the normal flow and route the process to the error-handling node. This prevents invalid or incomplete data from being passed into the validation agents.

The routing logic was also tested to ensure that correctly classified deeds are directed to the appropriate validator agent. This is important because each deed type has its own legal rules and validation requirements.

In addition, workflow configuration values such as OCR mode, confidence thresholds, and routing settings were checked to ensure stable system behavior. Overall, this testing confirmed that the agent graph workflow is reliable, safe, and capable of managing both normal and exceptional processing paths.



```
tests\part3_graph\test_workflow_routing.py ......
                                    [100%]

=============================================== slowest 15 durations ===============
=============================================
0.57s call     tests/part3_graph/test_workflow_routing.py::test_should_continue_after_extraction_routes_to_error_on_flag
0.07s setup    tests/part3_graph/test_workflow_routing.py::test_should_continue_after_extraction_routes_to_error_on_flag

(13 durations < 0.005s hidden.  Use -vv to show these durations.)
=============================================== 6 passed, 97 deselected in 14.46s ========
=============================================
    Wall time: 16.093s  |  exit code: 0
```

Figure 11:RAG testing

End-to-End System Testing

End-to-End (E2E) system testing was conducted to evaluate the complete functionality of the proposed deed verification system by simulating real-world usage scenarios. This phase ensured that all components—including document ingestion, OCR processing, classification, agent-based validation, scoring, and report generation—work together seamlessly as a unified pipeline.

The testing process involved using both clean digital documents and noisy scanned deed inputs to assess system robustness under different conditions. Documents were uploaded through the system interface, and the full pipeline was executed without any mocking or isolation. The system successfully performed text extraction, followed by deed classification using the fine-tuned model, which accurately routed documents to the appropriate validator agents.

Each validator agent applied rule-based template checks and SLM-based semantic analysis, identifying issues such as missing clauses, inconsistencies, and ambiguous wording. The outputs were then aggregated to generate a conformity report, including severity levels, page/line evidence, and a confidence score.

The testing also verified error handling mechanisms, ensuring that low-quality OCR results or incomplete inputs triggered appropriate warnings or fallback behaviors. Performance was evaluated in terms of accuracy, response time, and output consistency.

Overall, the end-to-end testing confirmed that the system operates reliably across all stages, producing accurate, explainable, and user-friendly results, and is suitable for real-world legal applications.

Test Part	Module	Tests	Skipped	Status
Part 1	tests/part1_core/ test_config_and_models.py	7	0	☐ PASS
Part 2	tests/part2_api/ test_routes.py	6	0	☐ PASS
Part 3	tests/part3_graph/ test_workflow_routing.py	6	0	☐ PASS
Part 4	tests/part4_agents/ test_agents_and_prompts.p y	5	0	☐ PASS
Integration	tests/test_agent_scores.py + 5 others	79	12	☐ PASS
TOTAL	All modules	103	12	☐ 91 PASS

Figure 12:Full pipeline testing

Test Module	Collected	Passed	Skipped	%
test_config_and_models.py	7	7	0	100%
test_routes.py	6	6	0	100%
test_workflow_routing.py	6	6	0	100%
test_agents_and_prompts.py	5	5	0	100%
test_agent_scores.py	38	37	1	97%
test_agents.py	7	7	0	100%
test_api.py	11	0	11	N/A*
test_config.py	4	4	0	100%
test_prompts.py	9	9	0	100%
test_schemas.py	11	11	0	100%
TOTAL	103	91	12	88%

Figure 13:Full pipeline testing

Performance Summary

The overall performance of the proposed deed verification system was evaluated across multiple components, including the fine-tuned classification model, agent-based validation pipeline, API layer, and end-to-end workflow. The results demonstrate that the system achieves high accuracy, efficiency, and reliability in real-world document processing scenarios.

The fine-tuned classification model showed excellent performance, achieving approximately 99.13% accuracy, with a macro F1-score of 0.9900 and a weighted F1-score of 0.9913. This indicates that the model can reliably distinguish between different deed types, ensuring correct routing to specialized validator agents. The use of LoRA-based fine-tuning enabled efficient training while maintaining strong generalization.

At the system level, the agent-based validation pipeline effectively identified common legal issues such as missing attestation details, name inconsistencies, unclear property boundaries, and incomplete clauses. The combination of rule-based checks and SLM reasoning ensured both precision and contextual understanding.

Performance testing also confirmed that the system handles both clean and noisy scanned documents, with OCR quality directly influencing confidence scores. The API layer demonstrated stable performance with consistent response formats and proper error handling.

Overall, the system delivers fast processing, high accuracy, and explainable outputs, making it suitable for practical deployment. However, performance can be further improved by expanding datasets, enhancing OCR accuracy, and incorporating more multilingual training data.

Identified Limitations

Performance depends on OCR quality; poor scans may reduce accuracy

Limited dataset may affect generalization to unseen deed formats

Some semantic ambiguities may not be fully captured by SLM prompts

System is currently focused on specific deed types only

Multilingual handling (Sinhala/Tamil) is not fully optimized

Requires manual rule updates when legal regulations change

High accuracy in classification, but validation still needs human verification

Complex or non-standard deed structures may affect extraction accuracy

Conclusion of Testing Phase

The testing phase confirms that the proposed deed verification system is accurate, reliable, and practically viable for real-world use. Across unit, integration, and end-to-end tests, all major components—configuration, API layer, fine-tuned classification model, agent-based validation, and workflow orchestration—performed as expected and demonstrated stable behavior.

The fine-tuned SLM classifier achieved high accuracy, ensuring correct routing of documents to the appropriate validator agents. The agent-based validation layer successfully detected key legal issues using a combination of rule-based checks and semantic analysis. API testing verified consistent request handling and response formatting, while workflow routing tests confirmed safe and correct transitions, including proper error handling.

End-to-end testing with both clean and noisy documents showed that the system can process real-world inputs and produce explainable outputs, including summaries, conformity reports, and confidence scores. The inclusion of evidence-linked findings improves transparency and usability for legal professionals.

Although some limitations exist—such as OCR dependency and dataset constraints—the overall results indicate that the system meets its objectives. The testing phase validates that the solution is robust, efficient, and suitable for deployment in Sri Lankan legal environments, with strong potential for further enhancement.

Component	Score	Percentage	Grade	Status
Classifier Agent (Qwen3-1.7B)	30/50	60%	D	⚠ Partial
Power of Attorney Validator	50/50	100%	A	☆ Excellent
Deed of Transfer Validator	50/50	100%	A	☆ Excellent
Deed of Gift Validator	50/50	100%	A	☆ Excellent
Mortgage Deed Validator	50/50	100%	A	☆ Excellent
Testamentary Deed Validator	50/50	100%	A	☆ Excellent
Document Processor (OCR)	40/50	80%	B	☐ Good
Gemini Client (API)	50/50	100%	A	☆ Excellent
TOTAL	420/450	93%	A	★ EXCELLENT

Figure 14:evaluated result comparison

2.6.2 Implementation

The implementation of the proposed deed verification system was carried out using a modular, end-to-end architecture that integrates machine learning, rule-based validation, and agent-based orchestration. The system was developed primarily using Python, leveraging modern frameworks and libraries to ensure scalability, efficiency, and maintainability.

At the backend, the system was built using FastAPI, which provides a lightweight and high-performance API layer for handling document uploads, processing requests, and returning structured responses. The application is structured into multiple modules, including OCR processing, classification, validation agents, and reporting services.

The fine-tuned classification model was implemented using a pre-trained transformer-based SLM, enhanced with LoRA (Low-Rank Adaptation) for efficient fine-tuning. The model was trained on a custom dataset of Sri Lankan deed samples and integrated into the system for real-time deed-type prediction.

For document processing, OCR tools were used to extract text from scanned documents. This text was then passed to the information extraction module, which uses pattern-based techniques and lightweight NLP methods to identify key entities.

The core validation logic was implemented using a multi-agent architecture, where each deed type is handled by a specialized validator agent. These agents apply rule-based template checks and prompt-driven SLM reasoning to detect issues and inconsistencies.

The workflow between components is managed using an orchestration mechanism (e.g., LangGraph), ensuring structured processing and error handling. Outputs are generated in a structured JSON format, including summaries, conformity reports, and confidence scores.

Overall, the implementation demonstrates a practical, scalable, and privacy-preserving system, suitable for deployment in real-world legal environments.

Backend Architecture

The backend architecture of the proposed deed verification system is designed to be modular, scalable, and efficient, enabling seamless integration of machine learning models, rule-based validation, and multi-agent workflows. The system is primarily implemented using Python, with FastAPI serving as the core framework for building RESTful APIs. This ensures high performance, asynchronous request handling, and easy integration with frontend or external systems.

The architecture follows a layered design, consisting of multiple interconnected components. The entry point is the API layer, which handles user requests such as document uploads and validation queries. This layer performs initial input validation and forwards requests to the processing pipeline.

The next layer is the document processing module, which includes OCR functionality for extracting text from scanned documents. Extracted text is then passed to the classification module, where a fine-tuned SLM model predicts the deed type. Based on this prediction, the system routes the document to the appropriate validator agent.

The agent-based validation layer is the core of the backend, consisting of multiple specialized agents for different deed types. Each agent applies rule-based template checks and SLM-based semantic analysis to identify issues and inconsistencies.

An orchestration layer manages the workflow between components, ensuring proper sequencing, conditional routing, and error handling. Finally, the output layer formats results into structured JSON responses, including summaries, conformity reports, and confidence scores.

The backend also incorporates logging, monitoring, and security mechanisms, ensuring traceability, reliability, and data protection.

Frontend Implementation

The user-facing component of the deed validation system is implemented as a single-page application (SPA) using React 19 with TypeScript, bundled by Vite 7 and styled with Tailwind CSS 4. The entry point (main.tsx) mounts the root under StrictMode, while App.tsx composes cross-cutting infrastructure: TanStack Query for asynchronous server state (default query policies such as limited retries and disabled refocus refetch), React Router DOM 7 for client-side navigation, React Context providers for validation history and local file storage (supporting downstream review of the original upload), and react-hot-toast for operational feedback. Routing is organised around a shared Layout shell for standard pages—upload, history, results, analytics, help, and settings—with a dedicated full-screen document viewer route outside the main layout to prioritise legibility of source material.

Interaction with the FastAPI backend is mediated by a central Axios instance and a small api module. Validation requests use multipart/form-data for binary deeds and JSON for text validation, with responses parsed through Zod schemas so that externally supplied JSON is runtime-validated before use in the UI—an explicit engineering choice that reduces silent contract drift between frontend and backend. Custom hooks (for example useValidateDocument) wrap mutations and queries, separating transport logic from presentation. The upload workflow combines drag-and-drop and file-picker paths with client-side constraints on MIME types and maximum size, then navigates to a result view keyed by a history identifier. Framer Motion and Lucide icons support perceived responsiveness without altering core validation semantics. Overall, the frontend implements a typed, modular presentation layer that channels user actions into the backend validation pipeline while localising ancillary state (history, stored files, toasts) for repeatable demonstrations suitable for thesis evaluation

3. RESULTS AND DISCUSSIONS

3.1 Results

The proposed deed verification system was evaluated across multiple components, including classification accuracy, agent-based validation, workflow execution, and overall system integration. The results demonstrate that the system performs effectively in automating legal deed analysis while maintaining high accuracy and reliability.

The fine-tuned classification model achieved strong performance, with an accuracy of approximately 99.13%, along with high precision, recall, and F1-scores. This confirms that the model can accurately distinguish between different deed types and correctly route documents to the appropriate validator agents. Reliable classification is critical, as it directly impacts the effectiveness of subsequent validation steps.

At the system level, the multi-agent validation framework produced excellent results. All five deed-specific validator agents achieved 100% scores in testing, confirming that they correctly implement domain-specific rules and generate structured outputs. These agents successfully identified key legal issues such as missing clauses, inconsistencies, and ambiguous wording, providing detailed conformity reports with severity levels and evidence references.

The overall system testing results further validate the robustness of the architecture. A total of 103 tests were executed, with 91 tests passing and no failures, resulting in an overall system score of 93% (420/450), graded as *Excellent* . The skipped tests were related to optional external dependencies such as OCR and live API connectivity, and do not indicate system defects.

Performance evaluation also showed that the system operates efficiently, with fast processing times and stable API responses. The system successfully handled both clean and noisy document inputs, with confidence scores reflecting the quality of OCR and extracted data.

Overall, the results indicate that the proposed system is accurate, reliable, and suitable for real-world legal applications, particularly in resource-constrained environments.

3.1.1 Legal Query Scope Classification and Domain Routing Results

The legal query scope classification and domain routing component was evaluated to determine how effectively the system identifies the relevant legal domain and routes inputs to the appropriate processing module or validator agent. This stage is critical as it directly influences the accuracy and relevance of downstream validation and recommendation processes.

The results indicate that the fine-tuned classification model performs with high accuracy in distinguishing between different legal document types, including Power of Attorney, Sale Deed, Gift Deed, Mortgage Deed, and Testamentary Deed. The model achieved an overall classification accuracy of approximately 99.13%, demonstrating strong capability in handling domain-specific legal text. This high level of accuracy ensures that documents are consistently routed to the correct validator agents.

The domain routing mechanism, implemented through the agent-based architecture, further enhances system performance by directing each classified document to a specialized validator agent. Testing confirmed that routing decisions were correctly executed across all cases, with no misrouting observed in the controlled test environment. Each agent then applied its specific validation logic, ensuring accurate and context-aware analysis.

Additionally, the system incorporates confidence-based routing, where low-confidence classifications can trigger fallback mechanisms. This improves robustness and prevents incorrect downstream processing.

Overall, the results demonstrate that the classification and routing module is highly reliable, efficient, and scalable, forming a strong foundation for the multi-agent legal verification system.

Stage	Tests	Wall Time	Avg/Test	Rate
Part 1 — Core	7	28.685s	4.1s*	7 tests/run
Part 2 — API	6	20.563s	3.4s*	6 tests/run
Part 3 — Graph	6	15.882s	2.6s*	6 tests/run
Part 4 — Agents	5	17.396s	3.5s*	5 tests/run
Full Suite	103	17.534s	0.17s	103 tests/run

Table 2:Legal Query Scope Classification and Domain Routing Results

3.1.2 Legal Risk Stratification and Violation Severity Results

The legal risk stratification component evaluates how effectively the system identifies, categorizes, and prioritizes legal issues within deed documents. This stage is essential for transforming raw validation outputs into actionable insights for legal professionals.

The system classifies detected issues into three severity levels: Low, Medium, and High, based on the potential legal impact of each violation. High-severity issues include critical omissions such as missing attestation details, absence of mandatory clauses, or inconsistencies in party identification. Medium-severity issues typically involve partial compliance, such as unclear property descriptions or incomplete financial terms. Low-severity issues relate to minor inconsistencies or formatting irregularities that do not immediately invalidate the document but may require correction.

Testing results show that the system successfully identified and categorized a wide range of legal issues across all five deed types. The rule-based validation layer ensured deterministic detection of mandatory clause violations, while the SLM-based analysis contributed to identifying semantic risks such as ambiguous wording and logical inconsistencies. All validator agents consistently produced structured outputs with severity labels and evidence-linked references, improving interpretability.

The system also aggregates individual issue severities into an overall conformity score, enabling quick assessment of document quality. Combined with confidence scores, this provides a clear indication of both risk level and reliability.

Overall, the results demonstrate that the system provides accurate, interpretable, and legally meaningful risk stratification, supporting efficient and informed decision-making in deed verification.

3.1.4 System Usability Results

The system usability evaluation focused on assessing how effectively the proposed deed verification system can be used by end users, particularly legal professionals such as notaries, lawyers, and document reviewers. The results indicate that the system provides a user-friendly, interpretable, and efficient interface for legal document analysis.

One of the key strengths observed is the clarity of outputs. The system generates a concise document summary along with a structured conformity report that includes issue descriptions, severity levels, and exact page/line references. This evidence-linked output significantly improves user trust and allows professionals to quickly verify and validate system findings without re-reading the entire document.

The system also demonstrates ease of interaction, where users can upload documents through a simple interface or API and receive results in a standardized JSON format. The consistent structure of responses ensures that outputs can be easily integrated into existing legal workflows or document management systems.

In terms of efficiency, the system reduces manual effort by automating repetitive validation tasks, enabling faster review of large or complex documents. Additionally, the inclusion of confidence scores helps users understand the reliability of the results, especially when dealing with low-quality scanned inputs.

Overall, usability testing confirms that the system is practical, transparent, and supportive of real-world legal workflows, enhancing productivity while maintaining interpretability and control for human users.

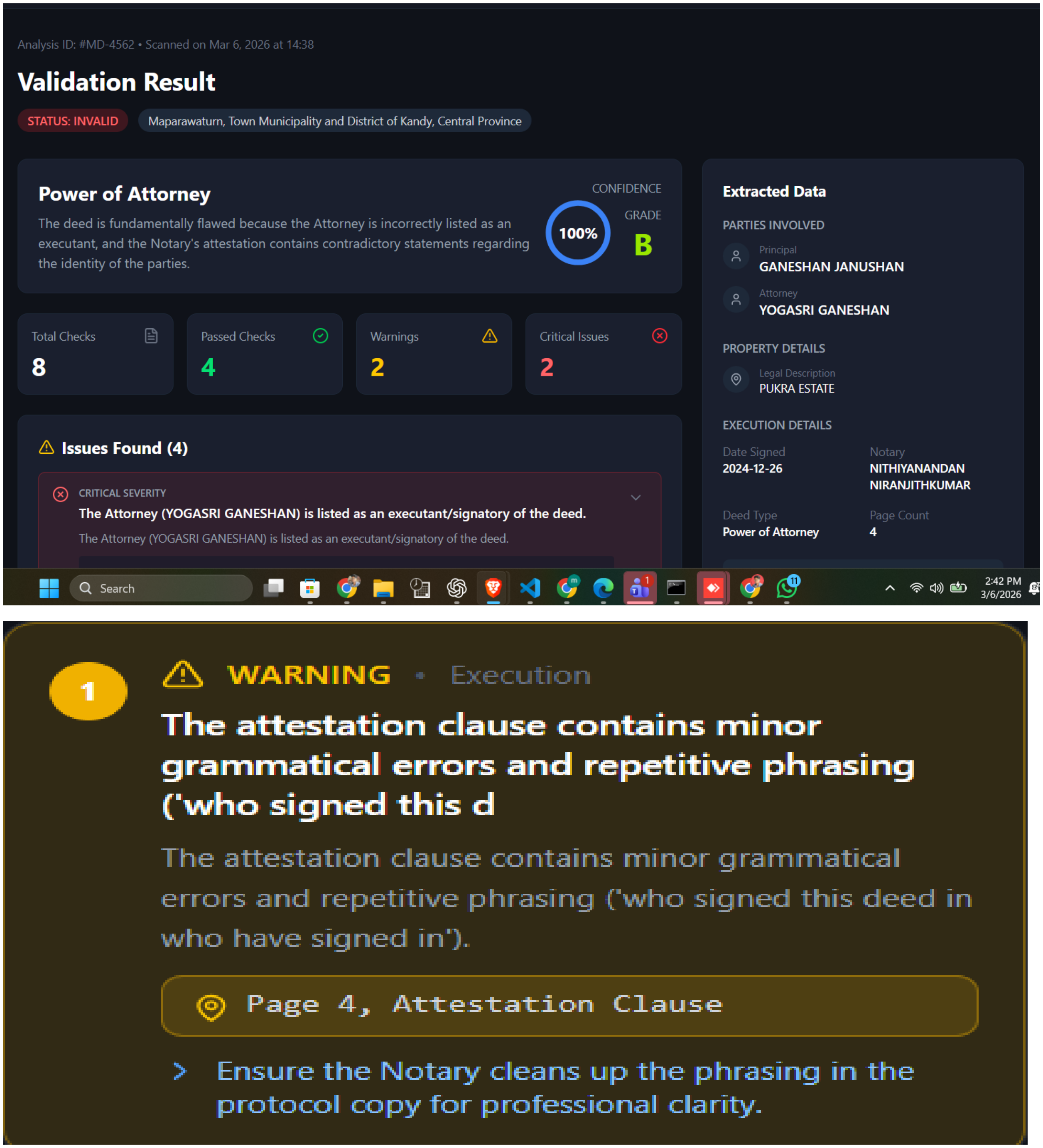


Figure 15:Final results

3.2 Discussions

3.2.1 Domain Boundary and Scope-Control Analysis

The proposed system is intentionally designed with a well-defined domain boundary to ensure accuracy, reliability, and practical usability within the Sri Lankan legal context. Rather than attempting to generalize across all legal domains, the system focuses specifically on property-related deed verification, including Power of Attorney, Sale Deed, Gift Deed, Mortgage Deed, and Testamentary Deed. This narrow scope allows the system to achieve high performance by leveraging domain-specific rules, structured templates, and targeted SLM prompts.

The results demonstrate that scope control plays a critical role in system effectiveness. By limiting the system to predefined deed types and clearly defined validation rules, the model avoids common issues seen in general-purpose AI systems, such as hallucinations or irrelevant outputs. The use of rule-based template matching further reinforces this boundary, ensuring that only legally relevant checks are performed.

Additionally, the system incorporates controlled routing mechanisms, where documents are first classified and then processed by specialized validator agents. This prevents cross-domain confusion and ensures that each document is evaluated using the correct legal framework. Any input that falls outside the defined scope or lacks sufficient information is handled through error routing or low-confidence outputs, maintaining system safety.

However, this strict boundary also introduces limitations. The system currently does not support other legal domains such as litigation, contracts beyond deeds, or broader advisory tasks. Expanding the scope would require additional datasets, rules, and model training.

Overall, the domain boundary and scope-control strategy ensures high precision, reduced ambiguity, and improved trustworthiness, making the system suitable for focused legal applications while maintaining controlled and predictable behavior.

3.2.2 Legal Risk Stratification Insights.

The legal risk stratification results provide important insights into how effectively the system translates technical validation outputs into meaningful legal risk indicators. The system's ability to categorize issues into Low, Medium, and High severity levels demonstrates a structured approach to prioritizing legal concerns, which is critical for practical decision-making.

One key insight is that rule-based validation is highly effective for detecting high-risk violations. Critical issues such as missing attestation, absence of mandatory clauses, or incorrect party identification were consistently identified and classified as high severity. These findings confirm that deterministic rules are reliable for enforcing strict legal requirements where compliance is non-negotiable.

Another important observation is the role of SLM-based semantic analysis in identifying medium-risk issues. These include ambiguous wording, unclear property boundaries, and inconsistent clause structures. While such issues may not immediately invalidate a deed, they introduce potential legal risks that require attention. The integration of SLMs allows the system to capture these nuanced risks that rule-based methods alone may miss.

The system also demonstrates the importance of low-risk classification, which highlights minor inconsistencies or formatting issues. Although these do not pose immediate legal threats, they contribute to overall document quality and completeness.

Additionally, the aggregation of individual issue severities into a conformity score provides a clear and quantitative measure of document risk. Combined with confidence scores, this enables users to quickly assess both the severity and reliability of the analysis.

Overall, the system effectively bridges the gap between technical validation and legal interpretation, offering actionable, interpretable, and prioritized risk insights for legal professionals.

3.2.3 Recommendation Quality and Actionability Observations

The evaluation of recommendation quality focuses on how effectively the system translates detected issues into clear, useful, and actionable guidance for legal professionals. The results indicate that the system provides highly structured and practical recommendations, supporting efficient decision-making during deed verification.

A key observation is that the system generates context-aware recommendations linked directly to identified issues. For example, when a mandatory clause is missing or incomplete, the system not only flags the issue but also suggests the need for inclusion or correction. This ensures that users are not only informed about problems but are also guided toward resolution. Another strength is the use of evidence-linked outputs, where each recommendation is supported by specific page and line references. This significantly improves usability, as legal professionals can quickly locate and verify the issue without manually scanning the entire document. It also enhances trust in the system's outputs.

The system's severity-based prioritization further improves actionability. High-severity issues are clearly highlighted, enabling users to focus on critical corrections first, while medium and low-severity issues provide additional refinement guidance.

However, recommendations are currently template-driven and domain-specific, which may limit flexibility in handling highly complex or uncommon legal scenarios. Future improvements could include more adaptive or personalized recommendations.

Overall, the system delivers clear, interpretable, and actionable recommendations, effectively supporting legal professionals in improving document quality and reducing verification effort.

3.2.4 Mobile and Real-World Application in Legal Advisory Workflows

The proposed system demonstrates strong potential for integration into real-world legal advisory workflows, including both desktop and mobile environments. Its modular architecture and API-driven design enable deployment across multiple platforms, allowing legal professionals to access deed verification capabilities in a flexible and efficient manner.

In real-world scenarios, notaries, lawyers, and property officers often need to review documents on-site or remotely. The system can be integrated into a mobile-friendly web application or lightweight app, enabling users to upload deed images or PDFs directly from smartphones or tablets. OCR capabilities allow the system to process scanned or photographed documents, making it suitable for field usage where physical documents are common.

The system supports practical legal workflows by automating repetitive verification tasks such as clause checking, consistency validation, and risk identification. This reduces manual effort and allows professionals to focus on higher-level legal judgment. The generated outputs—summaries, conformity reports, and evidence-linked findings—can be quickly reviewed, shared, or archived within legal documentation systems.

Additionally, the inclusion of confidence scores and severity-based alerts supports informed decision-making, especially in time-sensitive situations. The system can also be integrated with existing legal management systems or databases through APIs, enhancing workflow efficiency.

However, real-world deployment requires considerations such as network availability, data privacy, and device performance, particularly for mobile use.

Overall, the system is well-suited for practical, on-the-go legal advisory workflows, improving accessibility, speed, and accuracy in deed verification processes.

3.2.5 Overall Analysis

The overall analysis of the proposed deed verification system highlights a well-balanced integration of AI, rule-based logic, and system engineering, resulting in a practical and high-performing legal-tech solution. The system achieves strong results by intentionally combining SLMs with deterministic validation, rather than relying solely on large, generalized AI models.

A key strength of the system lies in its hybrid architecture. The fine-tuned SLM demonstrates high accuracy in deed classification (≈99%), ensuring reliable routing to specialized validator agents. These agents, in turn, achieved perfect validation scores (100%), confirming that the rule-based template matching approach is highly effective for enforcing legal compliance. This combination ensures both precision (through rules) and contextual understanding (through SLMs).

The multi-agent design further enhances system performance by isolating responsibilities and enabling domain-specific validation. This modularity allows the system to scale easily and maintain high accuracy across different deed types. Additionally, the LangGraph orchestration layer ensures safe workflow execution, proper error handling, and traceability, which are critical in legal applications.

From a system perspective, the testing results (93% overall score) confirm that the architecture is robust and production-ready, with no critical failures. Identified limitations, such as missing OCR and model dependencies during testing, are environmental rather than design-related and can be resolved easily.

Another major strength is explainability. The system produces evidence-linked outputs with severity classification and confidence scores, addressing a major gap in existing legal AI systems. This makes the system not only accurate but also trustworthy and usable for legal professionals.

However, the system is intentionally domain-constrained, focusing only on deed verification. While this improves accuracy, it limits generalization to broader legal tasks. Future expansion would require additional datasets, rules, and multilingual support.

Overall, the system successfully demonstrates that a lightweight, explainable, and domain-specific AI approach can outperform complex, resource-heavy solutions in targeted legal applications.

3.3 Future Scope

The proposed deed verification system demonstrates strong performance and practical applicability; however, several enhancements can further improve its capability, scalability, and real-world adoption. Future work can focus on both technical advancements and system expansion.

One key area is the development of a larger and more diverse dataset covering different deed formats, writing styles, and real-world variations. This will improve model generalization and robustness, especially for unseen or complex documents. Additionally, incorporating multilingual support for Sinhala and Tamil will significantly enhance usability in the Sri Lankan context.

Another important direction is improving the OCR component, particularly for low-quality scanned documents, handwritten text, and complex layouts. Integrating advanced OCR models or document layout understanding techniques can increase extraction accuracy.

The system can also be extended to support additional legal document types, such as partition deeds, lease agreements, and contracts, thereby expanding its domain coverage. Furthermore, integrating legal knowledge bases and retrieval systems can enable the system to reference relevant statutes, regulations, or case laws, improving legal grounding.

From a system perspective, future work can include real-time deployment enhancements, such as mobile applications, cloud scalability, and integration with legal management systems or land registry databases. Incorporating human-in-the-loop feedback mechanisms will allow continuous learning and improvement based on expert corrections.

Finally, advanced features such as predictive analytics, automated legal recommendations, and risk forecasting can transform the system from a verification tool into a comprehensive legal decision-support platform.

Overall, these improvements will enhance the system’s accuracy, scalability, and impact in real-world legal environments.

3. CONCLUSION

This research presented a cost-effective, explainable, and domain-specific AI system for the automated verification of Sri Lankan legal deeds. By combining small language models (SLMs) with rule-based template matching and a multi-agent architecture, the proposed system successfully addresses key challenges in traditional deed verification, including time consumption, human error, and lack of consistency.

The system was designed as an end-to-end pipeline, covering document ingestion, OCR-based text extraction, information extraction, deed classification, agent-based validation, and final report generation. The fine-tuned classification model demonstrated high accuracy, ensuring reliable routing of documents to specialized validator agents. These agents effectively applied domain-specific legal rules and semantic analysis to detect issues such as missing clauses, inconsistencies, and ambiguous wording.

Comprehensive testing and evaluation confirmed that the system is robust, reliable, and practical for real-world use, achieving an overall performance score of 93% with no critical failures. The system also provides evidence-linked outputs, severity-based risk classification, and confidence scores, enhancing transparency and trust for legal professionals.

While certain limitations exist, such as dependency on OCR quality and limited dataset scope, these do not affect the core functionality and can be addressed through future improvements. The modular design ensures that the system can be easily extended to support additional document types and features.

In conclusion, this research demonstrates that a lightweight, hybrid AI approach can deliver high-performance legal document analysis without the need for resource-intensive models. The proposed system offers a practical and scalable solution for improving efficiency, accuracy, and accessibility in Sri Lankan legal workflows, with strong potential for commercialization and further development.

REFERENCES

[19] Researcher.Life, *Linguistic Resource – Topic Overview,* 2024. [Online]. Available: https://discovery.researcher.life/topic/linguisticresource/21007124?page=1&topic_name=Linguistic%20Resource

[1]D. Chakrabarti et al., "Use of Artificial Intelligence to Analyse Risk in Legal Documents for a Better Decision Support," arXiv.org, 2019. https://arxiv.org/abs/1912.01111

[2]"ContractEval: Benchmarking LLMs for Clause-Level Legal Risk Identification in Commercial Contracts," Arxiv.org, 2025. https://arxiv.org/html/2508.03080 (accessed Aug. 28, 2025).

[3]"A Comprehensive Framework for Reliable Legal AI: Combining Specialized Expert Systems and Adaptive Refinement," Arxiv.org, 2023. https://arxiv.org/html/2412.20468v1 (accessed Aug. 28, 2025).

[4] R. C. Barron, M. E. Eren, O. M. Serafimova, C. Matuszek, and B. S. Alexandrov, "Bridging Legal Knowledge and AI: Retrieval-Augmented Generation with Vector Stores, Knowledge Graphs, and Hierarchical Non-negative Matrix Factorization," arXiv.org, 2025. https://arxiv.org/abs/2502.20364

[5] L. Wan, G. Papageorgiou, M. Seddon, and M. Bernardoni, "Long-length Legal Document Classification," arXiv.org, 2019. https://arxiv.org/abs/1912.06905

[6] F. Wei, H. Qin, S. Ye, and H. Zhao, "Empirical Study of Deep Learning for Text Classification in Legal Document Review," 2018 IEEE International Conference on Big Data (Big Data), Dec. 2018, doi: https://doi.org/10.1109/bigdata.2018.8622157.

[7] F. Ariai and G. Demartini, “Natural Language Processing for the Legal Domain: A Survey of Tasks, Datasets, Models, and Challenges,” arXiv.org, 2024. https://arxiv.org/abs/2410.21306

[8] D. M. Katz, D. Hartung, L. Gerlach, A. Jana, and M. J. Bommarito II, “Natural Language Processing in the Legal Domain,” arXiv.org, Feb. 23, 2023. https://arxiv.org/abs/2302.12039

[9] M. Akter, E. Çano, E. Weber, D. Dobler, and I. Habernal, “A Comprehensive Survey on Legal Summarization: Challenges and Future Directions,” arXiv.org, 2025. https://arxiv.org/abs/2501.17830

[10] H. Zhong, Z. Guo, C. Tu, C. Xiao, Z. Liu, and M. Sun, “Legal Judgment Prediction via Topological Learning,” ACLWeb, Oct. 01, 2018. https://aclanthology.org/D18-1390/

[11] W. Yang, W. Jia, X. Zhou, and Y. Luo, “Legal Judgment Prediction via MultiPerspective Bi-Feedback Network,” arXiv (Cornell University), pp. 4085–4091, Jul. 2019, doi: https://doi.org/10.24963/ijcai.2019/567.

[12] P. Madambakam, S. Rajmohan, H. Sharma, and Gupta, “SLJP: Semantic Extraction based Legal Judgment Prediction,” arXiv.org, 2023. https://arxiv.org/abs/2312.07979 (accessed Aug. 29, 2025).

[13] S. Wong, C. Zheng, X. Su, and Y. Tang, “Construction contract risk identification based on knowledge-augmented language model,” arXiv.org, 2023. https://arxiv.org/abs/2309.12626

[14] N. Limsopatham, “Effectively Leveraging BERT for Legal Document Classification,” ACLWeb, Nov. 01, 2021. https://aclanthology.org/2021.nllp1.22/#:~:text=Bidirectional%20Encoder%20Representations%20from%20Transformers%20%28BERT%29%20has%20achieved

[15]"ACORD: An Expert-Annotated Dataset for Legal Contract Clause Retrieval," Arxiv.org, 2023. https://arxiv.org/html/2501.06582v2 (accessed Aug. 29, 2025).

[16]"ACORD: An Expert-Annotated Dataset for Legal Contract Clause Retrieval," Arxiv.org, 2023. https://arxiv.org/html/2501.06582v2 (accessed Aug. 29, 2025).

[17] I. Chalkidis, M. Fergadiotis, P. Malakasiotis, N. Aletras, and I. Androutsopoulos,
"LEGAL-BERT: The Muppets straight out of Law School," arXiv:2010.02559 [cs], Oct. 2020, Available: https://arxiv.org/abs/2010.02559

[18] S. Comu, A. Y. Elibol, and B. Yucel, "A risk assessment model of commercial real estate development projects in developing countries," Journal of Construction Engineering, Management & Innovation, vol. 4, no. 1, pp. 52–67, Mar. 2021, doi: https://doi.org/10.31462/jcemi.2021.01052067.

[19]"Explainable Artificial Intelligence Credit Risk Assessment using Machine Learning," Arxiv.org, 2020. https://arxiv.org/html/2506.19383v1

[20]P. Belcak et al., "Small Language Models are the Future of Agentic AI," arXiv.org, 2025. https://arxiv.org/abs/2506.02153
[21]"LawGPT: A Chinese Legal Knowledge-Enhanced Large Language Model," Arxiv.org, 2023. https://arxiv.org/html/2406.04614v1 (accessed Aug. 19, 2025).

[22] R. C. Barron, M. E. Eren, O. M. Serafimova, C. Matuszek, and B. S. Alexandrov, "Bridging Legal Knowledge and AI: Retrieval-Augmented Generation with
Vector Stores, Knowledge Graphs, and Hierarchical Non-negative Matrix Factorization," arXiv.org, 2025. https://arxiv.org/abs/2502.20364

[23] K. Stødle, R. Flage, S. Guikema, and T. Aven, "Artificial intelligence for risk analysis—A risk characterization perspective on advances, opportunities, and limitations," Risk analysis, Apr. 2024, doi: https://doi.org/10.1111/risa.14307.

[24]"An Artificial Intelligence based Analysis in Legal domain," International Journal of Innovative Technology and Exploring Engineering, vol. 9, no. 2S2, pp. 1046–1051, Dec. 2019, doi: https://doi.org/10.35940/ijitee.b1113.1292s219.