MAKERERE
P.O Box 7062 Kampala Uganda
Email: principal.cis@mak.ac.ug
Website: www.cocis.mak.ac.ug

UNIVERSITY
Telephone: +256-414-545029
Fax: +256-41 532780

COLLEGE OF COMPUTING AND INFORMATICS TECHNOLOGY

CONVERSATIONAL GRAPH RETRIEVAL AUGMENTED GENERATION FOR LEGAL LAND ADVISORY SERVICES

By Group4

DEPARTMENT OF COMPUTER SCIENCE

SCHOOL OF COMPUTING AND INFORMATICS TECHNOLOGY

This work has been supported by
Makerere University COCIS RISE and MakRIF

A Project Report Submitted to the School of Computing and Informatics Technology, For the Study Leading to a Project Report in Partial Fulfillment of the Requirements for the Award of the Degree of Bachelor of Science in Computer Science Of Makerere University

Supervisor

Ggaliwango Marvin

School of Computing and Informatics Technology, Makerere University
ggaliwango.marvin@cit.ac.ug, +256754655524

June 2025

GROUP MEMBERSHIP

SNo	Names	Registration Number	Signature
1	Owiny-Dollo Marvin	22/U/3783/EVE	
2	Natukunda Phionah	22/U/3658/PS	
3	Magino Daniel	22/U/21768/PS	
4	Kaitesi Joan	22/U/3141/PS	

Declaration

It is hereby declared that:

1. The thesis submitted is our own original work while completing the degree at Makerere University.
2. The thesis does not contain material previously published or written by a third party, except where this is appropriately cited through full and accurate referencing.
3. The thesis does not contain material which has been accepted, or submitted, for any other degree or diploma at a university or other institution.
4. We have acknowledged all main sources of help.

STUDENT SIGNATURES

Signature: ____________________ **Date:** ____________________
Name: Owiny-Dollo Marvin
Reg No: 22/U/3783/EVE
Student No: 2200703783

Signature: ____________________ **Date:** ____________________
Name: Natukunda Phionah
Reg No: 22/U/3658/PS
Student No: 2200703658

Signature: ____________________ **Date:** ____________________
Name: Magino Daniel
Reg No: 22/U/21768/PS
Student No: 2200721768

Signature: ____________________ **Date:** ____________________
Name: Kaitesi Joan
Reg No: 22/U/3141/PS
Student No: 2200703141

Acknowledgments

We begin by giving all glory and thanks to our Lord Jesus Christ, whose grace, strength, and guidance sustained us throughout the development and completion of this project.
We would also like to acknowledge the invaluable support provided by Makerere University through the COCIS RISE initiative and the MakRIF funding program, which significantly contributed to the successful execution of our work.
Our sincere appreciation goes to our supervisor, Ggaliwango Marvin, for his steadfast mentorship, insightful guidance, and encouragement throughout the project.
We are also grateful to everyone who supported us in various ways—through advice, motivation, or assistance on this journey.

List of Figures

List of Tables

Dedication

We would like to express our sincere gratitude to the following individuals for their invaluable support and guidance throughout this project. Our deepest thanks go to Ggaliwango Marvin, our supervisor, for his unwavering mentorship, expertise, and encouragement, which were instrumental in shaping and refining our ideas. We also extend our heartfelt appreciation to Dr. Rose Nakibule for her insightful contributions and guidance in defining what would constitute a final year project.
In addition, we acknowledge the support of other staff members, especially Mr. John Kizito, whose significant input during the internship training sessions greatly enhanced our understanding and approach.

Nomenclature

A list of symbols & abbreviations used within the body of the document

AI	Artificial Intelligence
BERT	Bidirectional Encoder Representations from Transformers
BLEU	Bilingual Evaluation Understudy
F1-Score	Harmonic mean of precision and recall
GPT	Generative Pre-trained Transformer
LSTM	Long Short-Term Memory
ML	Machine Learning
MSE	Mean Squared Error
NLP	Natural Language Processing
RAG	Retrieval-Augmented Generation
RNN	Recurrent Neural Network
ROUGE	Recall-Oriented Understudy for Gisting Evaluation
Seq2Seq	Sequence-to-Sequence Model
T5	Text-to-Text Transfer Transformer
TF-IDF	Term Frequency-Inverse Document Frequency
Transformer	Transformer Neural Network

Table of Contents

Abstract

Accessing accurate and comprehensible legal land information in Uganda remains a major challenge, particularly for individuals in rural areas who may face barriers such as limited awareness of legal procedures or lack of access to qualified legal professionals. The problem lies in effectively retrieving and generating context aware legal guidance from diverse, unstructured, and often scanned sources while ensuring that the information is interpretable, linguistically accessible, and robust against document quality issues. An AI powered land advisory system was developed that leverages Graph Retrieval Augmented Generation (Graph RAG) and Computer Vision to provide intelligent, context aware legal support. Graph RAG builds upon Retrieval Augmented Generation (RAG) by introducing a structured knowledge graph composed of entities, legal clauses, and inter document relationships, providing more interpretable responses across multiple legal sources, making it especially suitable for complex legal queries that require structured understanding. Computer Vision techniques were applied exclusively to scanned land title documents using deep learning models including CNN, ResNet, VGG16, VGG19, and Vision Transformer (ViT). An ensemble learning approach combining these models via majority voting significantly improved classification robustness, particularly in handling noisy scans, occlusions, or overlapping features. Among the individual models, ViT achieved the highest classification accuracy at 98%, while the ensemble achieved an average F1 score of 97% across all test samples. The semantic search component was implemented using SentenceTransformers and FAISS indexing, while generative question answering was handled by a fine tuned T5 model, trained over 50 epochs, achieving a validation accuracy of 98% with a validation loss of 0.45. User interaction is supported via a multilingual web interface, where system responses are automatically translated to Luganda using voice synthesis at the output stage to improve accessibility. The system demonstrates a scalable, intelligent conversational legal decision support tool for delivering legal land information through natural conversation and offers a practical solution to bridge the information gap in land governance, empowering citizens to make informed decisions regarding land rights and documentation.

Keywords: Conversational AI, Natural Language Processing, Multi-Modeling, Legal Advisory Services, Knowledge Graph Retrieval Augmentation, Computer Vision, Large Language Models, Machine Translation

Introduction

1.1 Introduction

Legal Advisory Services in Uganda are severely challenged by widespread land disputes, significantly hindering economic development and individual well being [2]. This situation is particularly critical because land serves as an essential resource for housing, agriculture, and livelihoods, playing an undeniably pivotal role in both personal welfare and national growth [25]. Uganda's land tenure system is complex with multiple land tenure frameworks like freehold, leasehold, customary, and mailo tenure [25]. This complexity, along with widespread limited access to accurate legal knowledge, results in the prevalence of land disputes and persistent legal uncertainty throughout the country [17]. Land wrangles stand out as one of Uganda's foremost legal challenges, with police data clearly indicating a substantial increase in land related crimes and fraud cases, surging by 46% from 2023 to 2024 alone [17], [23]. A considerable portion of civil court cases are directly related to land matters, contributing significantly to an overwhelming backlog that paralyzes the judicial system, where unresolved disputes alone lock up an estimated UGX 85 trillion [16]. These delays not only prevent the timely administration of justice but also impose severe financial repercussions on affected individuals [16]. Furthermore many citizens unknowingly acquire land with forged land titles or violate legal procedures due to a critical lack of awareness regarding proper land acquisition and transfer processes [23]. Formal legal services provided by lawyers, brokers, or district land boards are often prohibitively expensive or practically inaccessible, particularly for vulnerable rural populations who are most in need of such support [2]. These intermediaries are frequently associated with pervasive corruption and exploitative practices as well further diminishing public trust in official channels. Consequently, individuals often resort to obtaining informal advice from friends, family and unverified intermediaries, significantly increasing their risk of receiving misinformation and facing subsequent legal vulnerability [2]. This highlights an urgent and pressing need for innovative solutions capable of bridging the gap between legal complexity and public comprehension, especially within a context where digital literacy varies considerably across the population. Forged land titles are a critical issue in land administration, directly causing disputes and undermining property security throughout Uganda [23], [25]. Scanned or photographed land title documents are susceptible to forgery [7]. This inherent vulnerability shows the need for automated systems to accurately classify and extract information for efficient provision of legal advisory services [7]. Text based Natural Language Processing methods prove insufficient for effectively processing such visually compromised documents unlike Optical Character Recognition (OCR) which can analyze visual layout and verify document authenticity of land title documents [23]. While artificial intelligence (AI) can potentially address legal information gaps, general purpose conversational support tools lack the domain specificity required to accurately interpret the nuanced intricacies of Ugandan land laws or reliably verify official documents [12], [26]. Their broad design inherently prevents deep engagement with the specific nuances of Ugandan legal frameworks, which are absolutely crucial for providing accurate advice [19].

Natural Language Processing (NLP) is fundamental for the Conversational Intelligent legal advisory support tool to understand complex and diverse user queries in natural language [13], [25]. It enables the tool to parse unstructured legal texts, extract relevant information, and generate accurate answers [24], [26]. Given Uganda's multifaceted land tenure systems, NLP models are specifically fine tuned to capture jurisdiction specific terminology and document structures [25]. This precision enhances the providing of legally compliant advice and ensures users receive contextually appropriate responses [25].

Conversational AI leverages these NLP capabilities to simulate human like interactions, allowing users to engage in fluid, intuitive dialogue [7], [19]. By building on natural language understanding, conversational AI transforms static legal knowledge into an accessible and interactive interface. This continuous, multilingual tool extends legal empowerment, enabling users to clarify complex land issues in real time without the need for costly human experts [2], [6], [13].

Knowledge Graph Retrieval Augmentation enhances retrieval by organizing legal information into a structured graph of interconnected legal clauses, cases, and hierarchies [20], [22]. This semantic network allows the tool to perform precise, explainable reasoning across multiple sources, linking answers to relevant legal entities and facts [24], [26]. It builds upon the capabilities of conversational AI by grounding responses in authoritative knowledge and increasing transparency, which fosters user trust and legal literacy [2], [13], [25].

Large Language Models (LLMs) add generative power by producing coherent and contextually accurate legal responses [24], [26]. Fine tuned on land law datasets, these models synthesize information retrieved via knowledge graphs and semantic search, ensuring that generated answers are expert level and legally sound [12]–[14]. This integration supports the conversational AI in providing nuanced and comprehensive guidance to users [6].

Multi Modeling further extends the tool's abilities by intelligently combining both textual and visual data analysis [7]. Land document verification requires examining both the textual content and visual features like seals, signatures, and formatting [25]. Using ensemble learning of vision models such as CNN, ResNet, VGG16, VGG19, and Vision Transformer, the system validates document authenticity [7]. This multi modal approach improves the robustness of legal verification processes by intelligently linking visual evidence with textual law [26].

Computer Vision focuses specifically on detecting forgery and inconsistencies in scanned land certificates of title [7]. By analyzing features, the vision models can identify tampering and ensure the integrity of submitted land titles [25]. This visual verification works alongside NLP and multi-modeling techniques to cross validate legal documents, reducing errors and fraudulent claims [7].

Machine Translation enhances inclusivity by translating and vocalizing Conversational Intelligent legal advisory support tool responses into Luganda and other local languages [26]. This voice based module removes linguistic barriers and ensures that users who are not fluent or literate in English can access legal advice [2]. It enhances the conversational AI support tool by providing an oral interface that expands accessibility and user comprehension in rural and marginalized communities [7], [13], [24].

1.2 Problem Statement

In Uganda, most legal land advisory services were offered through face to face consultations with legal professionals. These services were not only expensive and time consuming but also largely inaccessible to individuals living in rural areas. Additionally, people with lower levels of education often found it difficult to understand complex legal procedures and land related documentation, exposing them to potential exploitation. While conversational AI tools such as ChatGPT and Gemini held promise, they struggled to provide accurate and relevant responses specific to Uganda's legal land framework. This was due to an absence of specific fine tuning on Uganda's legal policies and insufficient mechanisms for incorporating up to date legal regulations. Consequently, such tools often delivered overly generic or misleading answers that were unfit for Ugandan legal advisory needs. Moreover, these tools did not possess the ability to retrieve specific Uganda based legal land documents or verify the authenticity of scanned land titles. This posed a major risk to users who often encountered forged or tampered documents during land transactions. Verifying land title authenticity remained a significant challenge, especially in fraud prone areas. Thus, there was a clear need for a locally aware, automated tool that combined legal knowledge retrieval with document authentication capabilities.

1.3 Implemented Solution

To address the aforementioned challenges, a conversational AI chatbot was developed and deployed. This tool was enhanced with Retrieval Augmented Generation (RAG) and integrated computer vision capabilities. The objective was to provide legally accurate, context specific responses tailored to Uganda's land ownership laws, while also detecting forged land documents.

Initially, a domain specific dataset was curated. This dataset comprised official land documents, procedures, and legal guidelines sourced from established and trusted Ugandan institutions. Utilizing these documents, a RAG pipeline was implemented. This pipeline combined dense vector retrieval (using FAISS and sentence transformers) with a fine tuned text generation model based on T5. This architecture enabled the chatbot to fetch and respond with relevant legal context based on the user's query, supporting both English and local phrasing patterns.

For fraud detection, a computer vision module was integrated. This module was trained on scanned images of both authentic and forged Ugandan land titles. The visual classifier utilized a ResNet 50 model, which was fine tuned on annotated title documents. This classifier was capable of identifying visual discrepancies such as irregular stamps, altered fonts, and inconsistent seals. Users could upload scanned documents, and the tool would return a classification verdict (authentic or forged) with accompanying visual highlights.

This implementation successfully addressed the limitations of generic conversational tools by offering both accurate legal guidance and proactive fraud detection. It empowered users, especially those in rural or under served regions, to make informed land related decisions without exclusive reliance on legal professionals. By merging NLP and computer vision within a local context, this tool offered a robust, scalable solution for legal land advisory in Uganda.

1.4 Objectives

The following objectives guided the development and implementation of the Conversational AI legal advisory system tailored for Uganda's land sector:

1.4.1 Main Objective

To develop a scalable, accurate, and accessible Conversational AI system leveraging Retrieval-Augmented Generation (RAG) and computer vision to provide reliable legal land advisory services in Uganda.

1.4.2 Specific Objectives

1. To curate and pre process a comprehensive dataset of Uganda specific land related legal documents, ensuring accuracy, proper structure, and domain relevance.
2. To implement and evaluate advanced NLP techniques, including semantic search and Graph Retrieval-Augmented Fine-Tuning (GRAFT), for contextually relevant legal document retrieval.
3. To apply deep learning and ensemble computer vision techniques for accurate classification and verification of scanned Ugandan land title documents.
4. To design and deploy a conversational AI chatbot capable of understanding user intent, delivering accurate and context-aware legal advice, and supporting multilingual responses.
5. To deploy the system on a reliable server infrastructure with API access, ensuring broad accessibility and ease of integration with other platforms.

1.5 Research Questions

This research aimed to address the following key questions:

1. What legal land documents are required to support accurate, context aware responses in a Ugandan legal advisory system?
2. Which NLP techniques are most effective for retrieving legal documents relevant to user queries in the Ugandan context?
3. How can computer vision improve the classification and extraction of information from scanned Ugandan land title documents?
4. What design choices enable a conversational chatbot to deliver legal advice that is accurate, context aware, and multilingual?

1.6 Research Contributions

This project was able to contribute significantly to both natural language processing and legal technology domains through;

1. Providing an openly accessible, curated, and comprehensive dataset of Ugandan legal land documents and procedures, tailored specifically to Uganda's land administration context.

2. Delivering a Retrieval-Augmented Generation (RAG) powered conversational AI chatbot enhanced by Graph Retrieval-Augmented Fine-Tuning (GRAFT) that provides accurate, contextually relevant, and legally sound responses aligned with Uganda's unique land laws and practices.

3. Integrating advanced computer vision models to accurately classify and verify scanned land title documents, thereby improving the reliability of automated document handling in legal advisory services.

1.7 Scope

The scope of the project was focused on the design, implementation and deployment of an AI driven legal advisory system tailored to Uganda's land sector. The details are shown below

1. We curated and preprocessed Uganda specific legal land documents published before March 30, 2025. These were obtained from government sources including MLHUD, KCCA, NEMA, and ULC.

2. We Implemented semantic search and GRAFT techniques to enable precise and context aware retrieval of legal content relevant to user queries.

3. We Utilized computer vision methods, including deep learning and ensemble techniques, for the classification and verification of scanned land titles.

4. We developed a multilingual conversational chatbot capable of delivering accurate legal advice in English and Luganda.

5. We deployed the system on a robust server platform with API endpoints for broad accessibility and integration potential.

Related Works

2.1 Legal Land Advisory Systems

Navigating land matters in Uganda, including transactions, ownership verification, and dispute resolution, presents significant challenges such as high costs and rampant fraud [11], [25]. These issues are evident in several key areas. Firstly, individuals commonly visit physical offices to search land registers for ownership and property details, a process that is often time consuming and inefficient [15]. Secondly, engaging legal professionals for advice on land acquisition, transfer, or dispute resolution is common, but these services are prohibitively expensive for many, particularly rural and vulnerable populations [2]. Thirdly, land disputes, handled through courts or customary processes, are often protracted and costly, significantly contributing to court backlogs [16], with unresolved disputes alone locking up an estimated UGX 85 trillion in value [16]. The Ministry of Lands has introduced digital tools to digitize land records and streamline transactions for increased transparency and efficiency. However, these face several limitations currently. Many users, especially in rural communities, lack the necessary digital skills to effectively adopt and utilize these new platforms [2]. Furthermore, scanned or photographed land title documents, without robust security protocols, are susceptible to sophisticated forgery attempts, which undermines property security in the digital realm [7], [23]. Additionally, current digital tools that incorporate general purpose AI conspicuously lack the domain specificity required to accurately interpret Uganda's nuanced land laws or reliably verify official documents [12], [19], [26].

Conversational AI systems have been implemented and demonstrated across various domains to provide accessible, accurate, and real-time advisory support. These systems have successfully transformed static information into interactive interfaces. For instance, in accounting, conversational AI agents have been deployed to answer common queries about tax regulations or financial reporting standards, improving efficiency and accessibility for users [9]. By leveraging Natural Language Processing (NLP), deployed systems have effectively understood complex user queries, parsed unstructured texts, extracted relevant information, and generated accurate answers in diverse applications, including legal contexts [13], [25]. For specialized domains, NLP models have been successfully fine-tuned to capture domain-specific terminology and document structures, thereby enhancing contextually appropriate advice in existing implementations [25]. Knowledge Graph Retrieval Augmentation has been integrated into information systems to organize complex data into structured networks of interconnected entities and relationships [20], [22]. This semantic network has facilitated precise, explainable reasoning across multiple sources, allowing systems to link answers to authoritative entities and facts. This approach has proven effective in building trust and enhancing information reliability by grounding responses in verified knowledge within implemented solutions [24], [26]. Multi-Modeling capabilities have been implemented to intelligently combine both textual and visual data analysis for document verification and content understanding in various applications [7]. By integrating CNN and ResNet, systems were developed that can analyze visual features such as seals and signatures [7], [25], [26]. Large Language Models (LLMs), specifically when fine-tuned on specialized datasets, have produced coherent and contextually accurate responses by synthesizing retrieved information via knowledge graphs and semantic search [24], [26]. This has ensured expert-level and robust output from these systems in practice across various expert domains [12]–[14]. Machine Translation has been widely utilized in various digital applications to enhance inclusivity by translating and vocalizing responses into local languages. This voice-based module has removed linguistic barriers, ensuring users not fluent or literate in a primary language can access information, thereby expanding accessibility and user comprehension in diverse communities [7], [13], [24].

2.2 Conversational AI

Conversational AI enhances legal advisory systems through natural language interactions that align with human-computer interaction principles, offering intuitive and user-centric interfaces [19]. These systems handle diverse inputs and generate coherent responses, proving useful for automating tasks and improving workflows in land services [7], [26]. RAG integration ensures accurate and context-aware insights across legal frameworks [6]. Studies on chatbot evolution emphasize their adaptability to specific domains, though they struggle with ambiguity and incomplete queries [5]. Reinforcement learning-enhanced dialogue systems offer personalized interactivity but are computationally demanding [8]. Question-answering systems excel in parsing structured documents but falter with unstructured data common in land management [3]. RAG models bridge this gap by combining retrieval and generation, yielding nuanced responses but requiring robust document repositories [6]. Applications in urban planning demonstrate the efficacy of these systems in simplifying complex regulations and democratizing access to land services [7]. For instance, RAG-powered assistants can retrieve legal provisions and explain them in user-specific contexts, such as regional laws or document details [6]. This integration automates land records retrieval, dispute resolution, and legal verification—tasks traditionally marred by inefficiencies [7]. Especially in Uganda, such systems could alleviate burdens on legal institutions and make land services accessible to non-experts through real-time, tailored assistance [6].

2.3 Retrieval Augmentation

Retrieval Augmented Generation (RAG) systems have transformed legal land advisory services by enabling accurate and context specific responses through the integration of retrieval mechanisms and generative models [11]. A RAG system typically consists of a retriever, which identifies relevant documents, and a generator, which synthesizes responses based on retrieved content [13], [24]. Innovations like Long RAG accommodate longer contexts, while Graph RAG specifically enhances retrieval by structuring information as a knowledge graph, explicitly modeling entities and their complex relationships [20], [22].

Graph RAG builds upon traditional RAG by converting unstructured text into a network of interconnected nodes and edges, where nodes represent entities and edges define their relationships [22]. This graph based approach is particularly crucial for interpreting in-

terconnected legal clauses, statutes, and precedents, as it moves beyond simple semantic similarity to understand the logical connections between different pieces of information. It enables more precise and explainable retrieval through multi hop reasoning and semantic traversal, allowing the system to follow chains of relationships to find comprehensive answers that might be scattered across multiple documents [22]. For instance, in a legal context, GraphRAG can trace how a specific land transaction is affected by multiple regulations and specific land tenure provisions by navigating these relationships in the knowledge graph. This explicit representation of relationships significantly aids in grounding responses in a structured, verifiable knowledge base, thereby enhancing accuracy and reducing issues like hallucinations [22], [24].
Retrieval techniques such as Dense Passage Retrieval (DPR) and ColBERT further improve the accuracy and efficiency of document retrieval in complex domains, including highly specialized areas like Uganda's land law, encompassing Customary and Mailo systems [11], [24]. These systems dynamically generate legal advice tailored to regional regulations and document specific details [20]. By incorporating real time knowledge, RAG mitigates hallucinations and outdated responses, which is crucial for preventing fraud and interpreting evolving procedures [24], [25]. RAG systems also enhance access to services for marginalized communities by contextualizing complex frameworks and offering personalized guidance [26]. Their adaptability and scalability position them as key enablers of reliable information in fragmented environments like Uganda [13], [24].

2.4 Models and Techniques for Retrieval and Generation

Pretrained language models like GPT, BERT, T5, and BART form the foundation of Retrieval Augmented Generation (RAG) systems, providing coherent and contextually aware responses. When fine tuned for domain specific applications, such as legal land advisory services in Uganda, these models leverage general knowledge gained during pretraining to achieve superior performance across diverse language tasks [13], [14], [24], [26].
Comparative analyses consistently show that models like T5 and GPT2 outperform transformers trained from scratch. For instance, T5 Small (2% parameters of T5 Base) has surpassed baseline models in specialized tasks [13], [23]. Pretrained models demonstrated significantly higher accuracy on datasets requiring complex reasoning, such as the ListOps dataset where they achieved 58.7% accuracy versus 29.0% for scratch trained transformers [12], [14]. This highlights the profound impact of pretraining on a model's reasoning capabilities.
Despite their strengths, these models have inherent limitations, particularly when applied to highly specialized domains like land and legal advisory services in Uganda. A primary concern is their reliance on the quality and specificity of their pretraining data. If this data lacks sufficient legal or land related content pertaining to Ugandan law, the models may struggle to understand the nuanced terminology, historical context, or specific regulations unique to Ugandan land law. This can lead to inaccurate or incomplete advice if not thoroughly addressed during fine tuning [12], [26]. Another limitation is the computational cost and resource intensity associated with fine tuning and deploying these large models, which can be a barrier in resource constrained environments like those often found in Uganda.
Fine tuning remains essential for adapting these models to specific tasks. Even partial fine tuning (e.g., freezing attention layers) yields superior results compared to full training from scratch, though fully fine tuned models generally perform best [12], [13]. Domain

specific fine tuning allows these models to interpret and process complex content, such as intricate legal documents, by refining general reasoning into task specific logic [23], [24], [26]. This process is critical for ensuring the models can accurately understand and advise on specific land tenure systems, like Customary and Mailo, and their associated legal procedures relevant to Uganda.
By combining robust pretraining with strategic fine tuning, models like T5, BERT, BART, and GPT2 can effectively support dynamic and accurate retrieval and generation, even in nuanced domains like land management and legal advisory services, provided their domain specific limitations are carefully mitigated [12], [13].

2.5 Evaluation and Fine Tuning

Evaluating Retrieval Augmented Generation (RAG) systems involves several key metrics that quantify response quality, relevance, and ranking. These include Semantic Answer Similarity (SAS), ROUGE, BLEU, and Mean Reciprocal Rank (MRR) [1], [4], [10], [21]. Recall Oriented Understudy for Gisting Evaluation (ROUGE) assesses n gram, longest common subsequence (LCS), and skip bigram overlap, providing insights into both grammatical and semantic alignment [10], [25]. Bilingual Evaluation Understudy (BLEU) focuses on n gram precision and brevity, though it may sometimes underrepresent overall fluency [1]. Mean Reciprocal Rank (MRR) specifically emphasizes the position of the first relevant document retrieved, serving as a useful complement to broader metrics like precision, recall, and the F1 score [4]. These diverse retrieval and generation metrics are crucial for assessing both the retriever's contextual accuracy and the generator's fluency and coherence. For unlabeled data, faithfulness and relevance are key performance indicators, while accuracy is primarily emphasized in labeled datasets [1], [10].
Fine tuning is essential for enhancing model adaptation to specific domains, ensuring that RAG systems perform optimally in specialized contexts like Uganda's land and legal advisory services. Frameworks such as Hugging Face, Haystack, TensorFlow, and OpenAI API commonly support RAG system deployment and customization [12]. Advanced techniques, including low rank adaptation and various hybrid retrieval generation strategies, further refine model accuracy for highly specialized tasks [18]. RAG systems have demonstrated real world impact, notably in legal domains dealing with complex challenges like Uganda's land tenure. Here, tailored models are implemented to address data inconsistencies and ensure precise outputs, showcasing the practical utility of fine tuning [25]. Innovations like Graph RAG and Long RAG extend system capabilities through advanced graph based retrieval and robust long context handling, respectively [20], [22]. The synergy between document augmentation and strategic fine tuning further improves the efficiency and adaptability of these systems for enterprise use cases [26]. These continuous developments highlight the critical role of effective evaluation, strategic fine tuning, and tailored deployment in realizing the full potential of RAG systems across complex, data intensive domains.

Research Methodology

3.1 Multi Model Research Methodology.

The methodology below was used to implement a conversational AI chatbot, which was significantly enhanced with Graph Retrieval Augmented Generation (Graph RAG) and computer vision for detecting forged land certificates of title.

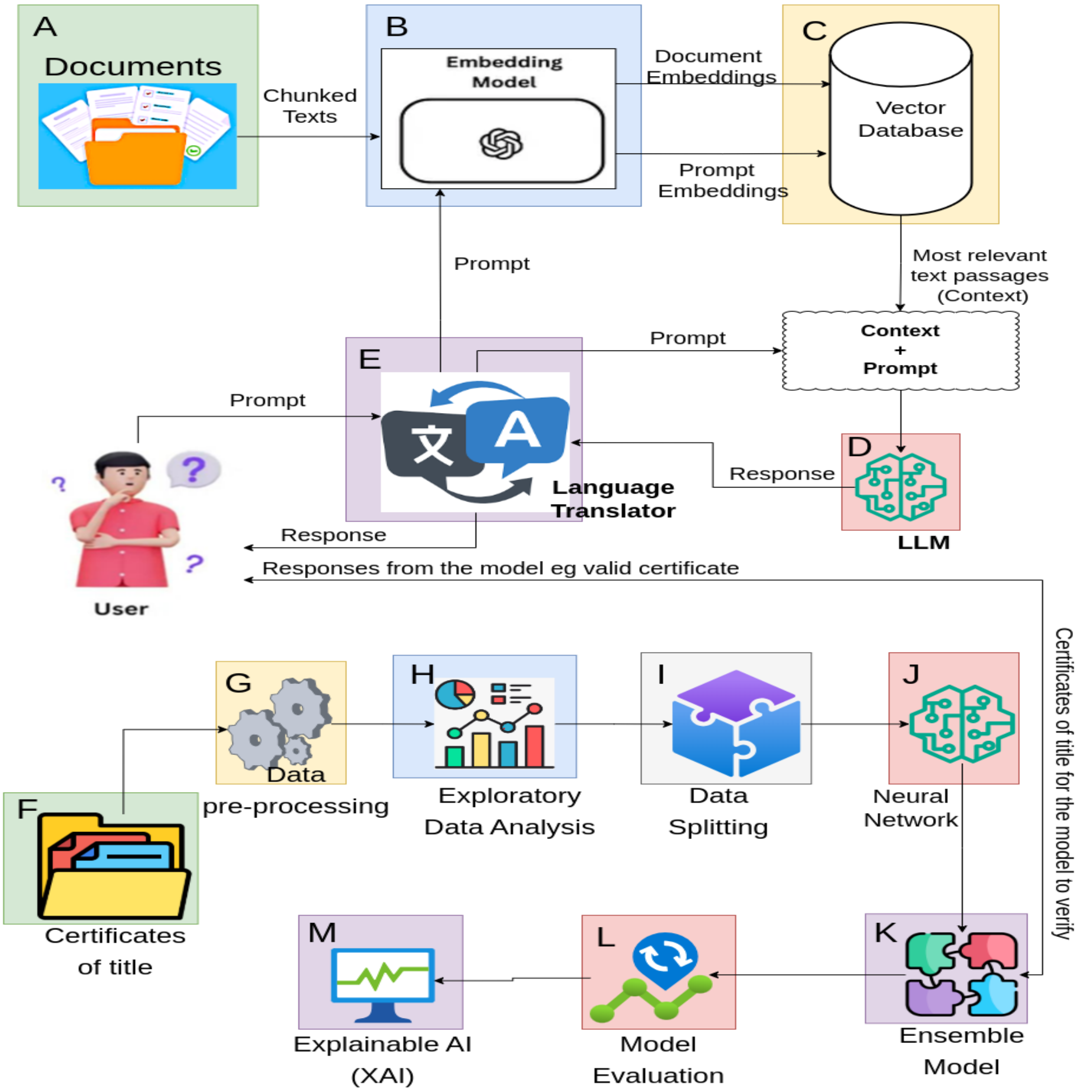

Figure 3.1: Research Methodology Diagram showing the implementation of the RAG enabled chatbot and the computer vision verification of the land certificates of title

A. Text Data Acquisition

The data for this study was sourced exclusively from reputable Ugandan governmental and public institutions involved in land related affairs. Legal documents (Figure 3.1, label A), were acquired through a combination of web scraping and direct collection from organizations such as the Ministry of Lands, Housing and Urban Development (MLHUD), Uganda Land Commission (ULC), Kampala Capital City Authority (KCCA), National Environment Management Authority (NEMA), and the Ugandan Parliament archives. This textual dataset was specifically curated for natural language processing (NLP) tasks.

B. Embedding Model

The cleaned textual data was transformed into vector embeddings using SentenceTransformers fine tuned on the legal corpus. These embeddings, generated by the model as shown in Figure 3.1 with label (B), captured semantic relationships between queries and legal content.

C. Vector Database

An Angora DB graph database (Figure 3.1, label C) was utilized to index these embeddings within a graph structure, enabling fast and accurate graph based retrieval and similarity search. Metadata tagging (e.g document source) was incorporated to enhance filtered search capabilities. This embedding and indexing pipeline formed the foundation of the system's graph based Retrieval Augmented Generation (RAG) capabilities.

D. LLM Model Integration

For each user query, the system retrieved relevant documents from the Angora DB graph database and passed them as context to the model. The Large Language Model (LLM) (Figure 3.1, label D) then utilized a graph RAG approach, responding based on its graph based legal data and real user queries to ensure contextual relevance and high quality responses. This setup allowed the chatbot to deliver accurate, jurisdiction specific legal guidance tailored to Uganda's land governance structure.

E. Language Translation

The Language Translator (Figure 3.1, label E) is crucial for enhancing the accessibility and user friendliness of the legal land information chatbot for local users in Uganda. Its primary function is to translate the generated responses from the Large Language Model (D) into Luganda. This ensures that users, particularly those who prefer or primarily communicate in Luganda, can receive legal guidance in a language familiar to them, thereby broadening the reach and utility of the system within diverse linguistic communities in Uganda.

F. Image Data Acquisition

Images of land certificates of title, specifically Freehold and Mailo titles (Figure 3.1, label F), were also acquired for the purpose of verifying their authenticity based on visual properties. These were acquired with variations in lighting, angles, resolution, and background clutter to reflect the challenges likely encountered during deployment.

G. Data Pre-processing

All collected images were pre processed to ensure consistency and readiness for downstream tasks, as shown in Figure 3.1 with label (G). Pixel normalization was done to scale pixel values to a uniform range, enhancing model stability and convergence during training. Optical Character Recognition (OCR) was then used to extract text from these images, which was subsequently subjected to a specific textual pre processing pipeline.

H. Exploratory Data Analysis (EDA)

Exploratory Data Analysis was conducted to uncover dominant patterns, topics and possible gaps within the corpus, as represented by Figure 3.1 with label (H). Frequency distributions and ngram (unigram, bigram, trigram) analysis was used to identify common legal phrases and terms. Visualizations such as bar charts and heatmaps highlighted thematic concentration across sources. This process also helped identify underrepresented topics in the data, guiding targeted data supplementation to improve domain coverage and retrieval accuracy.

I. Data Splitting

Following pre processing and EDA, the datasets were prepared for model training and evaluation through a strategic splitting process, denoted as Figure 3.1 with label (I). For this image dataset, a split of 80% for training, 10% for testing, and 10% for evaluation was applied to ensure reliable computer vision model performance.

J. Neural Network

Various Neural Network models were developed for the computer vision tasks, specifically for identifying document authenticity and flagging irregularities in scanned certificates of title. This is depicted (Figure 3.1, label J).

K. Ensemble Model

This phase involved the integration of the CNN and Vision Transformer (ViT) model to form a cohesive ensemble system. This overall integration is represented by Figure 3.1 with label (K).

L. Model Evaluation and Deployment

The Retrieval Augmented Generation (RAG) system was evaluated using a holistic framework focusing on real world performance dimensions beyond the metrics of accuracy, precision and recall. This evaluation process is broadly represented by Figure 3.1 with label (L).

M. Explainable AI (XAI)

The integration of Explainable AI (XAI) is depicted (Figure 3.1, label M). XAI aims to make the AI system's decisions and outputs transparent and understandable to users and developers. SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model agnostic Explanations) were used. SHAP provides a unified measure of feature importance by attributing the prediction to individual features, while LIME explains individual predictions by building a local interpretable model around the prediction. For a legal advisory system, interpretability is crucial for building trust and allowing users to verify the basis of the advice.

3.2 Corpus Description

The curated textual dataset consists of 41 publicly available legal and policy documents related to land governance in Uganda. These documents were sourced primarily from official publications, governmental repositories, and legal handbooks addressing various aspects of land law, policy, acquisition processes, land tenure systems, and registration procedures. The dataset includes acts such as the Uganda Land Act (1998), the National Land Policy, compulsory acquisition guidelines, land use and urban policy documents, and procedural guides for acquiring different land titles (freehold, leasehold, Mailo, and customary).

Characteristic	**Value**
Geographic scope	Uganda
Data collection period	Jan 2025 – March 2025
Document languages	English
File formats	PDF
Document types	Legislative acts, regulatory guidelines, land policy frameworks, procedural manuals, institutional reports
Total number of documents	41
Total word count (all docs)	75,430
Average words per document	1,839

Table 3.1: Comprehensive summary of the corpus characteristics, including document count, word statistics, file formats, languages, and collection details.

3.3 Exploratory Data Analysis (EDA)

The textual corpus was analyzed to identify its characteristics before further processing and modeling. The word cloud highlights the most frequently occurring terms, revealing dominant themes and keywords (Figure 3.2).

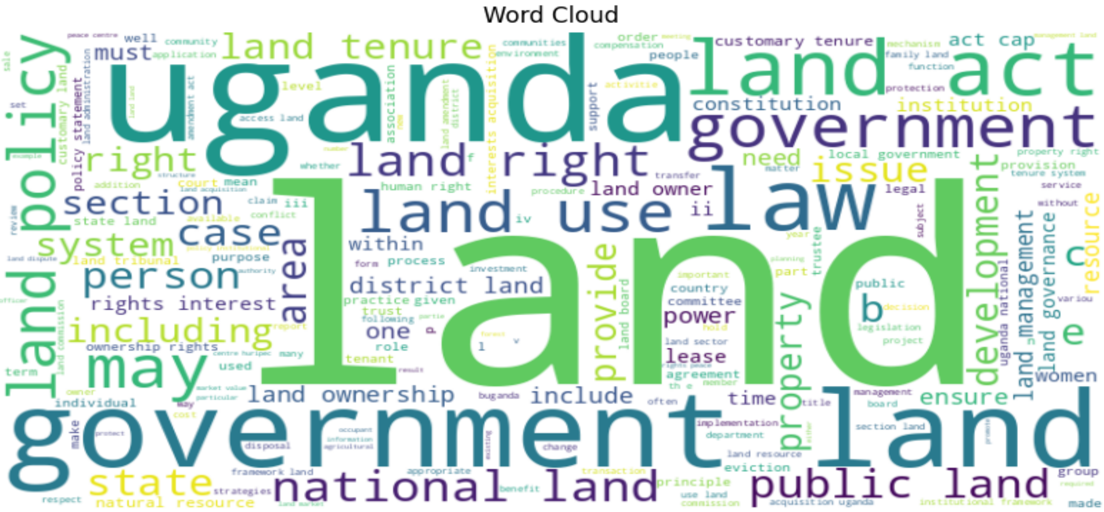

Figure 3.2: Word Cloud illustrating the most frequent terms in the corpus.

The unigram frequency distribution revealed the most common words in the corpus, with terms such as "land," "government," and "rights" appearing most frequently. These high-frequency terms aligned with the core themes of land governance and legal documentation.

The bigram frequency analysis highlighted frequently co-occurring word pairs, such as "government land" and "land act," reflecting specific and contextually relevant legal concepts within the dataset. This analysis provided deeper insights into phrase patterns prevalent in the documents.

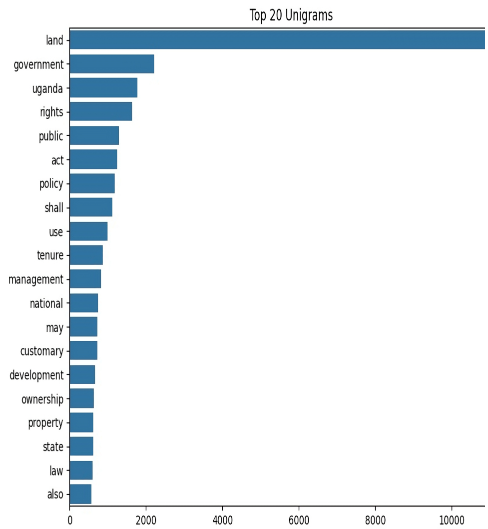

Figure 3.3: Top 20 Unigrams

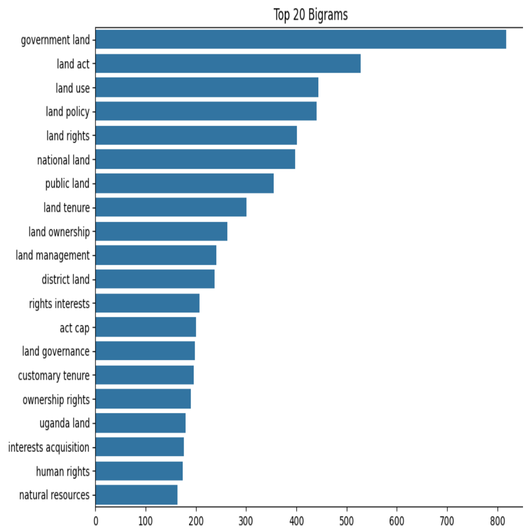

Figure 3.4: Top 20 Bigrams

The word count distribution before preprocessing reflects original document length variability with noise, while after preprocessing it shows a cleaner, standardized dataset with non-informative words removed for better analysis.

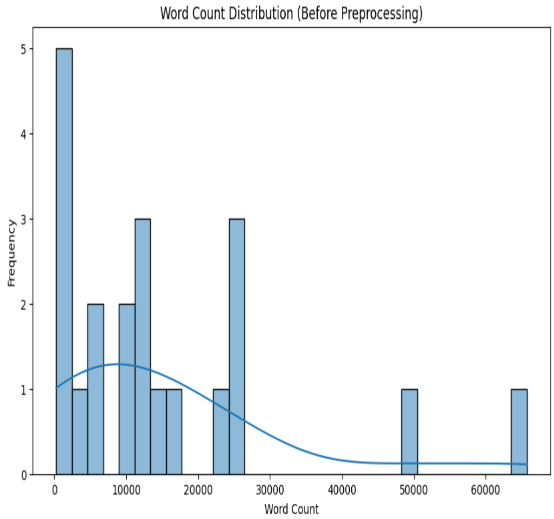

Figure 3.5: Word Count Distribution Before Preprocessing

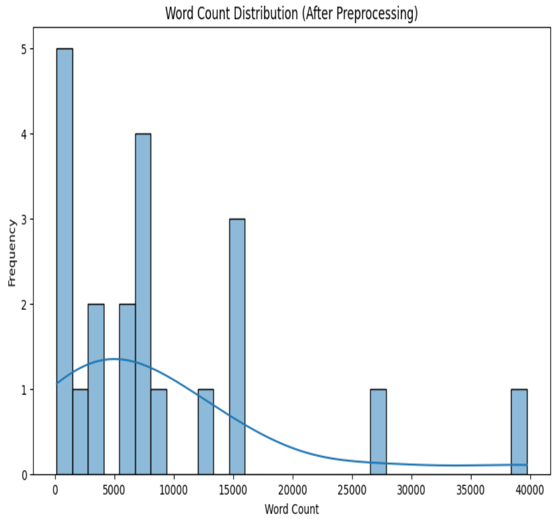

Figure 3.6: Word Count Distribution After Preprocessing

The box plot in Figure 3.7 shows the distribution of the top 20 word frequencies across the documents on a logarithmic scale, highlighting variability and outliers in word usage within the corpus.

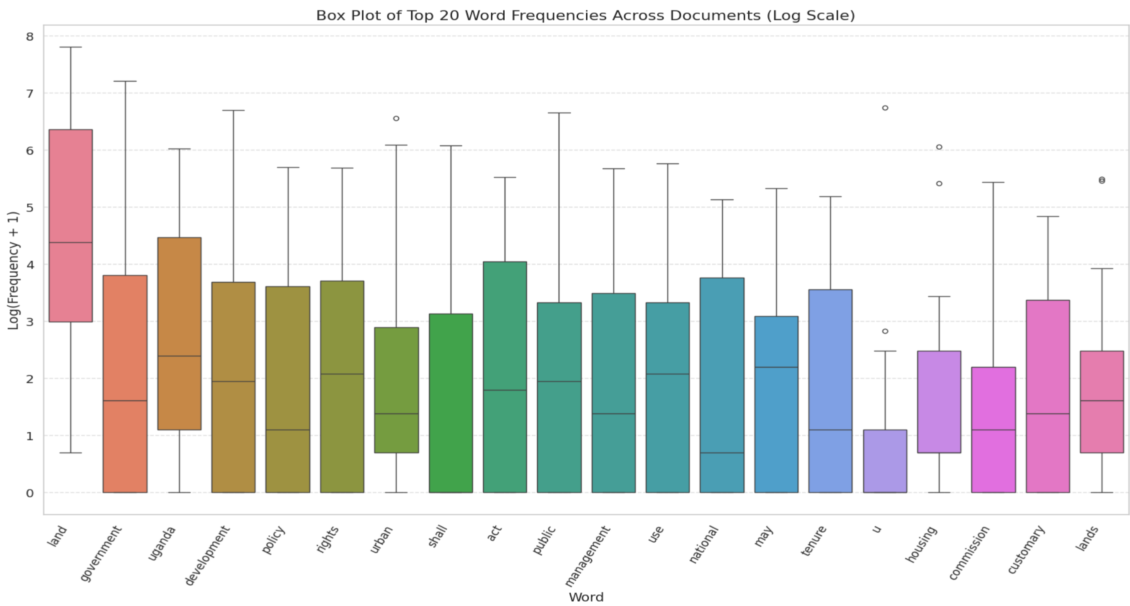

Figure 3.7: Box Plot of Top 20 Word Frequencies (Log Scale)

3.4 Model Evaluation

The selected evaluation questions shown in Table 3.2 focus on core aspects of land governance in Uganda and were carefully chosen for their relevance to both citizens and legal practitioners, as they reflect frequent information needs and practical challenges faced within the Ugandan land sector.

ID	Question Text
Q1	What are the four main types of land tenure systems recognized in Uganda?
Q2	How does one apply for a certificate of customary ownership in Uganda?
Q3	What is the role of the Ministry of Lands, Housing and Urban Development in land registration?
Q4	Can a foreigner own land in Uganda, and what are the restrictions, if any?
Q5	What steps are involved in converting leasehold land to freehold in Uganda?

Table 3.2: List of Questions Used in Model Evaluation

The comparative responses in Table 3.4 indicate that the created Graph chat tool consistently outperforms both DeepSeek and ChatGPT in delivering accurate and legally grounded answers to land related questions within the Ugandan context.

Question	DeepSeek	ChatGPT	Graph Chat tool
What are the four main types of land tenure systems recognized in Uganda?	Common types of land tenure globally include private ownership, leasehold, customary use, and government controlled land, though details vary by country.	The four types are customary, leasehold, freehold, and mailo. These systems determine how land is accessed, transferred, and inherited depending on legal frameworks and tradition.	According to the Land Act (Cap 227), Uganda recognizes four tenure systems: Customary (based on tradition), Leasehold (contractual and time-bound), Freehold (permanent private ownership), and Mailo (unique to Uganda, with overlapping rights and obligations).
How does one apply for a certificate of customary ownership in Uganda?	The process often involves filling an application with local land authorities, undergoing land inspection, and waiting for approval before receiving documentation.	In general, the applicant must approach local land committees, submit personal and land-related details, and await verification and processing by the land board.	Under Section 4 of the Land Act, an applicant submits a written request to the Area Land Committee. After site inspection and verification, the District Land Board, empowered by the Land Regulations (2004), may issue a Certificate of Customary Ownership (CCO).

Question	DeepSeek	ChatGPT	The Graph chat tool
What is the role of the Ministry of Lands, Housing and Urban Development in land registration?	Land ministries often oversee urban planning, surveying, land allocation, and title registration to ensure proper land governance.	The Ministry typically manages land-related policy, documentation, surveying, and issuing legal ownership titles. It plays a key role in urban development and planning.	The Ministry is responsible for policy formulation, oversight of the Uganda Land Information System (ULIS), and supervision of land title registration as guided by the Land Act and the Registration of Titles Act (Cap 230). It ensures compliance with land laws and supports decentralization through Zonal Offices.
Can a foreigner own land in Uganda, and what are the restrictions, if any?	Foreign nationals are usually subject to restrictions and may not own land outright in some countries, often having access through leases or partnerships.	Most countries place restrictions on foreign land ownership. Typically, foreigners can lease land for long periods but cannot acquire full ownership without specific approvals.	As per Article 237(2) of the 1995 Constitution and the Land Act, foreigners cannot hold land under freehold or mailo tenure. However, they may acquire land on leasehold basis for up to 99 years. Customary ownership is not available to non-citizens.
What steps are involved in converting leasehold land to freehold in Uganda?	Generally, one must apply to the relevant land board, show proof of lease and land use, and then follow procedures to gain approval for conversion.	The process includes submitting an application with supporting documents to the local land board or ministry, meeting eligibility criteria, and receiving confirmation after inspection.	In accordance with the Land Act and Land Regulations, a leaseholder submits an application to the District Land Board. After verifying compliance and inspecting the land, the Board recommends conversion. Upon approval, a Freehold Certificate is issued by the Ministry under the Registration of Titles Act.

Table 3.4: Comparative evaluation of responses from the general purpose LLMs (DeepSeek and ChatGPT) against the Graph chat tool on five land related legal questions. The responses are assessed based on their applicability and relevance within the Ugandan legal context.

DeepSeek produced generic responses applicable to a global audience while ChatGPT demonstrated moderate contextual awareness, with more detailed outputs but still lacking explicit references to Ugandan legal frameworks. The Graph chat tool demonstrated superior performance by integrating legal terminology and referencing statutes such as the Land Act (Cap 227) and the 1995 Constitution of Uganda. As shown in Table 3.4, this model provided the most contextually accurate, complete, and jurisdiction specific

responses, confirming its suitability for domain specific legal applications.
By evaluating model responses to these questions, we were able to assess factual accuracy of the responses produced.

3.5 Holistic Evaluation of the Created Tool

Holistic evaluation assessed the real world performance of LLMs beyond standard benchmarks. It was applied across summarization, information retrieval, bias, sentiment analysis, question answering, and toxicity detection to evaluate the strengths and limitations of each model within the context of Ugandan legal land information. This multi metric evaluation enabled a deeper understanding of not only how accurately the models generated content but also how ethically and responsibly they behaved in the sensitive land domain. Toxic or harmful outputs, for example, could have negatively impacted user trust. The comprehensive performance of the Deep, ChatGPT, and the Graph chat tool across six critical evaluation metrics is visually presented in Figure 3.8.

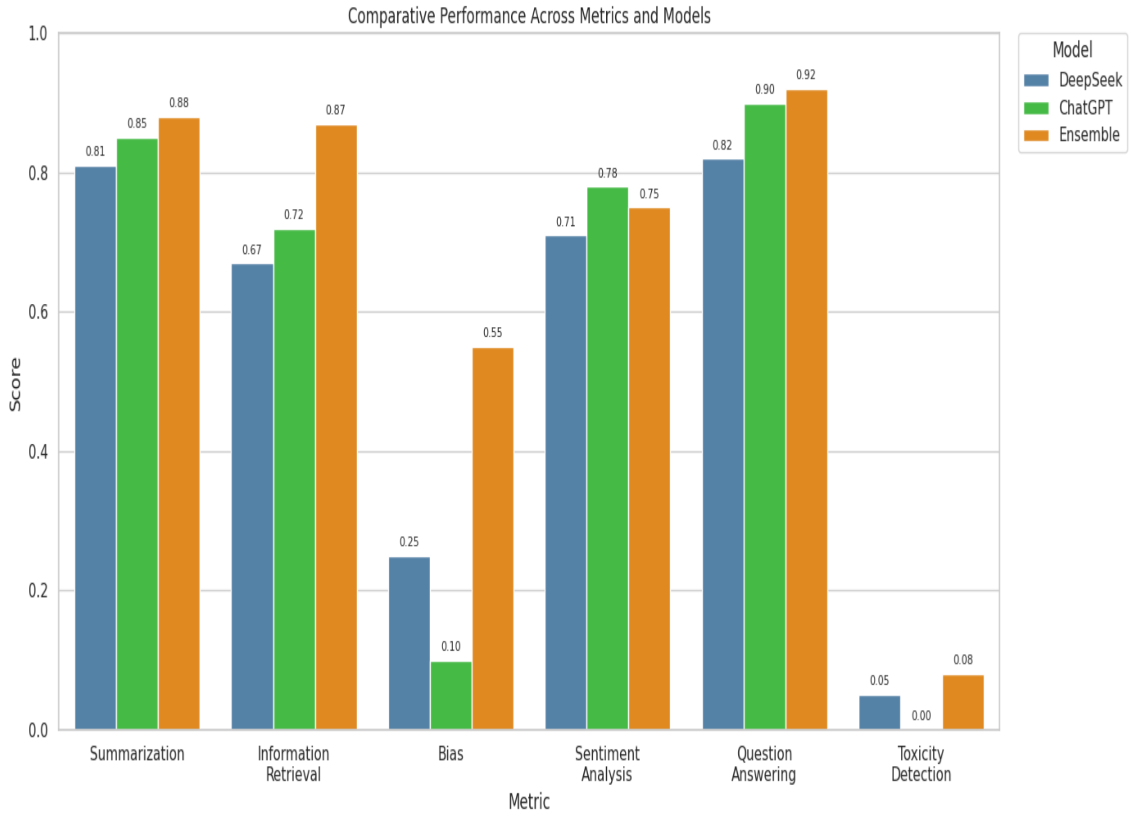

Figure 3.8: Comparative performance of the Deep, ChatGPT and the Graph chat tool across six key evaluation metrics. Lower scores indicate better performance for Bias and Toxicity Detection.

The evaluation in Figure 3.8 provided a holistic understanding of each model's capabilities within the context of providing legal land information. Overall, the Graph chat tool generally exhibited strong performance, particularly in tasks demanding precise information handling and accurate response generation.

Having reviewed the comprehensive overview of model performance, we'll now examine each holistic evaluation metric individually, exploring the specific insights each provides.

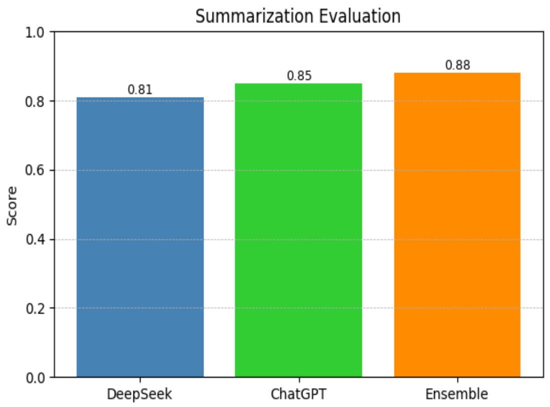

Figure 3.9: Summarization accuracy across models. The Graph chat tool scored highest at 0.88.

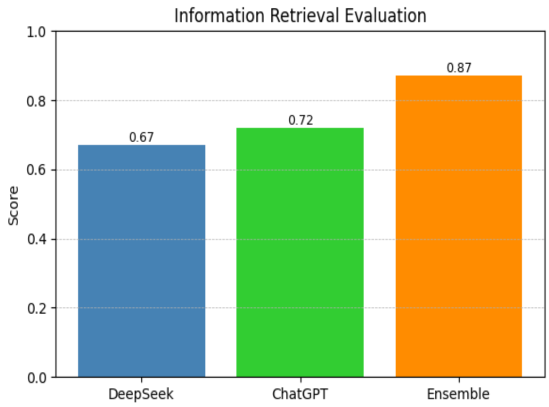

Figure 3.10: Retrieval performance across models. The Graph chat tool scored 0.87.

3.5.1 Summarization

Summarization evaluated how well a model could distill long legal responses into concise, relevant summaries. This was crucial for users with limited time or legal literacy. The evaluation in Figure 3.9 indicated strong summarization performance, with the Graph chat tool slightly outperforming the others.

3.5.2 Information Retrieval

Information retrieval (Figure 3.10) measured how accurately a model fetched relevant legal content in response to user queries. The Graph chat tool led with a score of 0.87, suggesting a superior ability to locate Ugandan land law details (Figure 3.10).

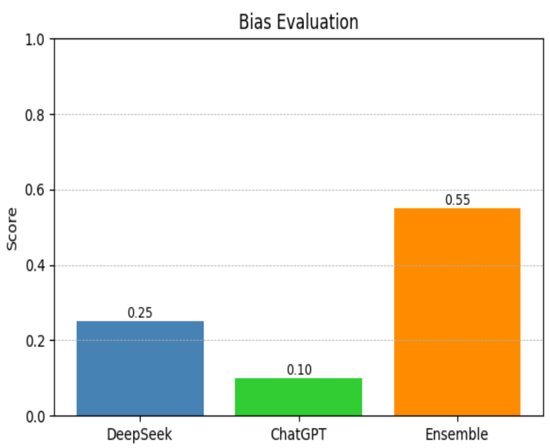

Figure 3.11: Bias scores based on gender neutral response evaluation. ChatGPT produced the most gender fair output, while the Graph chat tool showed moderate bias.

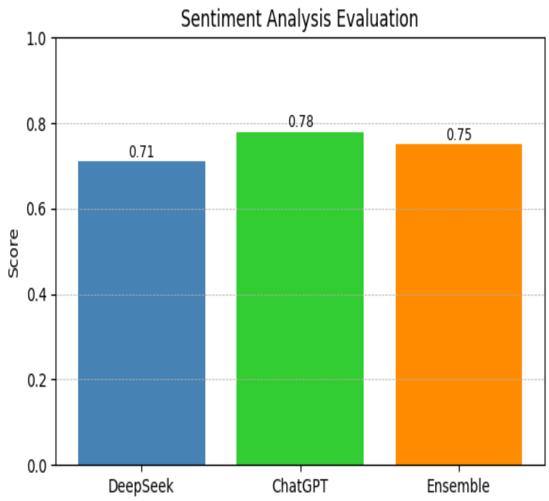

Figure 3.12: Sentiment neutrality comparison. ChatGPT offered the most balanced tone in legal land queries.

3.5.3 Bias

Bias was measured as preference or unequal referencing of gendered terms (e.g., “he” vs. “she”) in legal contexts. Lower scores indicated less gender bias. The visualization is shown in Figure 3.11.

3.5.4 Sentiment Analysis

Sentiment analysis gauged tone neutrality. For legal applications, objective and neutral sentiment was preferred. The results are shown in Figure 3.12.

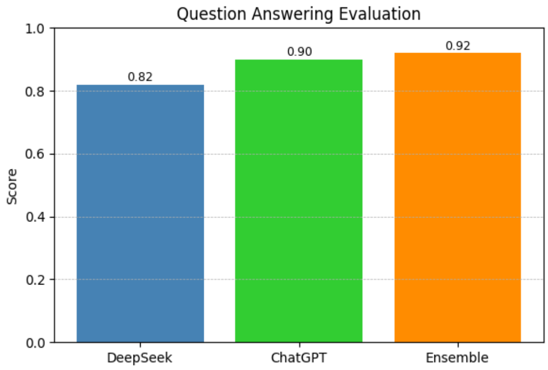

Figure 3.13: Question answering accuracy. The Graph chat tool provided the most legally accurate answers, closely followed by ChatGPT.

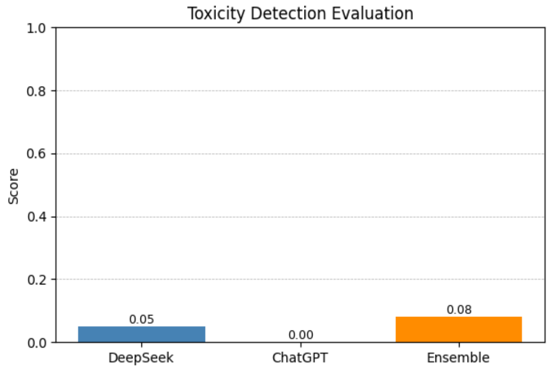

Figure 3.14: Toxicity scores across legal queries. All models showed non-toxic outputs, with ChatGPT scoring 0.0.

3.5.5 Question Answering

This metric assessed how accurately the model answered fact-based questions. It was central to legal assistant systems. The results are visualized in Figure 3.13.

3.5.6 Toxicity

Toxicity detection measured the presence of harmful or inappropriate language. All models remained well below the toxicity threshold, confirming the professionalism of their responses (Figure 3.14).

Metric	**Deep**	**ChatGPT**	**Graph Tool**
Summarization	0.81	0.85	0.88
Information Retrieval	0.67	0.72	0.87
Bias (Lower is Better)	0.25	0.10	0.55
Sentiment Analysis	0.71	0.78	0.75
Question Answering	0.82	0.90	0.92
Toxicity (Lower is Better)	0.05	0.00	0.08

Table 3.5: Performance of Models Across Evaluation Metrics

Table 3.6 presents a comparative overview of model-generated responses to five key land governance questions in Uganda, with each question accompanied by detailed answers in English and their respective translations in Luganda to enhance accessibility and contextual relevance for both language groups.

ID	Question Text	English Response	Luganda Response
Q1	What are the four main types of land tenure systems recognized in Uganda?	1. Mailo land tenure 2. Freehold land tenure 3. Customary land tenure 4. Leasehold land tenure These are established under Uganda's Constitution and regulated by the Land Act and the Registration of Titles Act.	1. Mailo land tenure 2. Freehold land tenure 3. Customary land tenure 4. Leasehold land tenure Zino zissibwawo wansi w'Etteeka ly'Ettaka n'Etteeka erikwata ku Kuwandiisa.
Q2	How does one apply for a certificate of customary ownership in Uganda?	1. Submit application to the Area Land Committee 2. Public notice is issued using Form 9 3. Hearing and verification are conducted 4. If approved, the Certificate of Customary Ownership is issued The process is governed by the Land Act and Land Regulations 2004. It protects vulnerable groups and formalizes customary rights.	1. Okuwaayo ebbaluwa y'okusaba eri Area Land Committee 2. Okutegeeza mu lusegere okuyita mu Form 9 3. Okukebera n'okuwuliriza 4. Okuweebwa Certificate of Customary Ownership ssinga kisalibwawo Enkola eno eyambako okukuuma abantu abatali bwenkanya era ewandiika mu mateeka obuyinza bw'obuwangwa.
Q3	What is the role of the Ministry of Lands, Housing and Urban Development in land registration?	1. Formulates land policy and laws 2. Oversees land registration 3. Issues land titles 4. Supports local land institutions 5. Links land registration with urban planning	1. Etongoza amateeka n'enkola ku ttaka 2. Efumintiriza n'okukulembera enkolagana y'amateeka 3. Ewa ttaayitulu z'ettaka 4. Ewagira amakomiti ag'ettaka mu masaza 5. Ekwataganya eby'okuwandiisa ettaka n'enkulaakulana y'ebibuga
Q4	Can a foreigner own land in Uganda, and what are the restrictions?	1. Foreigners can only acquire leasehold tenure for up to 99 years 2. Cannot own Mailo, Freehold, or Customary land 3. Lease must comply with Ugandan law 4. Ownership rights remain with the Ugandan lessor	1. Abatali bannansi basobola kufuna ettaka nga leasehold yokka okumala emyaka 99 2. Tebasobola kufuna Mailo, Freehold, oba Customary 3. Lease erina okugoberera amateeka ga Uganda 4. Ettaka liba lya nyini Uganda

ID	Question Text	English Response	Luganda Response
Q5	What steps are involved in converting leasehold land to freehold in Uganda?	1. Submit required forms and documents 2. Attach deed plans, title, photos, receipts, and DLO letter 3. If applicable, provide Surrender Deed 4. Present documents to Department of Land Administration 5. Land is verified and converted under the Land Act	1. Weereza formu ezikwetagisa n'empapula 2. Gattako deed plan, title, passport, receipts, ebbaluwa ya DLO 3. Weereza Surrender Deed ssinga kyetagisa 4. Weereza empapula mu Department of Land Administration 5. Ettaka likeberebwa ne likyusibwa nga bwekiri mu Land Act

Table 3.6: Models Responses to five land related questions in English translated to Luganda.

To ensure that the chatbot responses are both meaningful and reliable, several standard evaluation metrics were also applied.

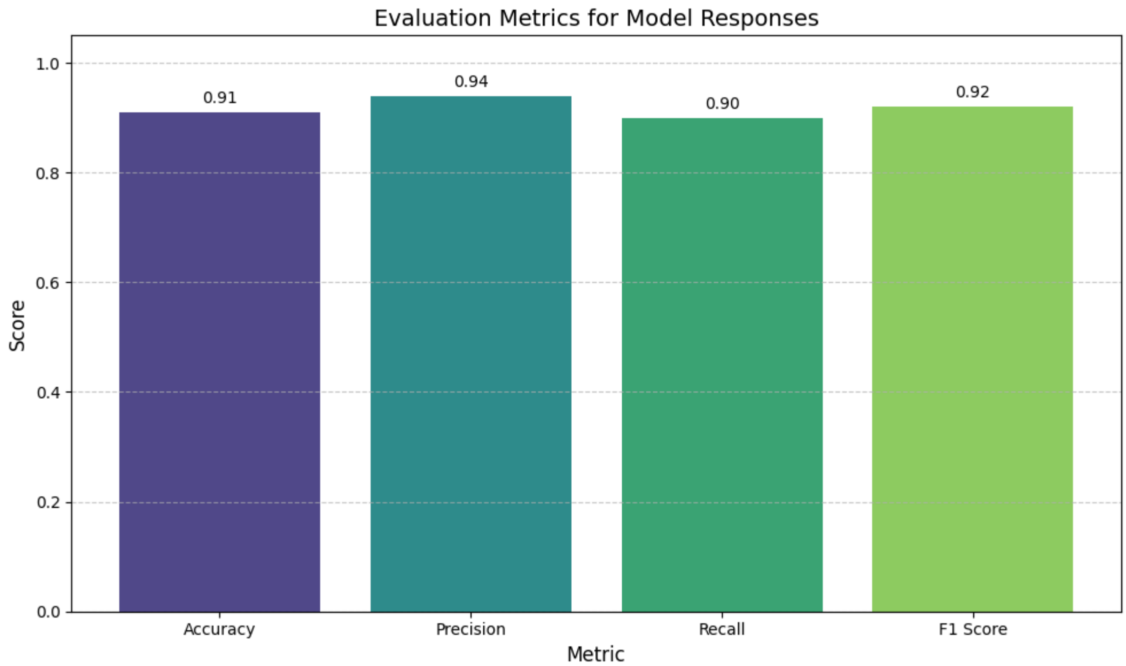

Figure 3.15: Evaluation results for Precision, Recall, Accuracy, and F1 Score. These metrics provided insights into how well the chatbot retrieves correct information and avoids incorrect or incomplete answers.

These metrics shown in Figure 3.15 assess various aspects of the conversational Advisory support tools' performance in terms of correctness, completeness, and consistency. These include Accuracy which quantifies the exact match proportion between conversational AI support tool responses and reference answers. Precision indicates the proportion of relevant responses among all responses generated by the conversational AI support tool. Recall evaluates the conversational AI support tool's ability to retrieve all relevant responses

compared to the reference answers. F1 Score, the harmonic mean of precision and recall, provides a balanced measure encompassing both false positives and false negatives.
ROUGE (Recall-Oriented Understudy for Gisting Evaluation) is a set of metrics commonly used to evaluate automatic summarization and machine translation systems. It compares the overlap of n-grams between the generated and reference responses. The higher the ROUGE score, the more similar the generated text is to the reference.

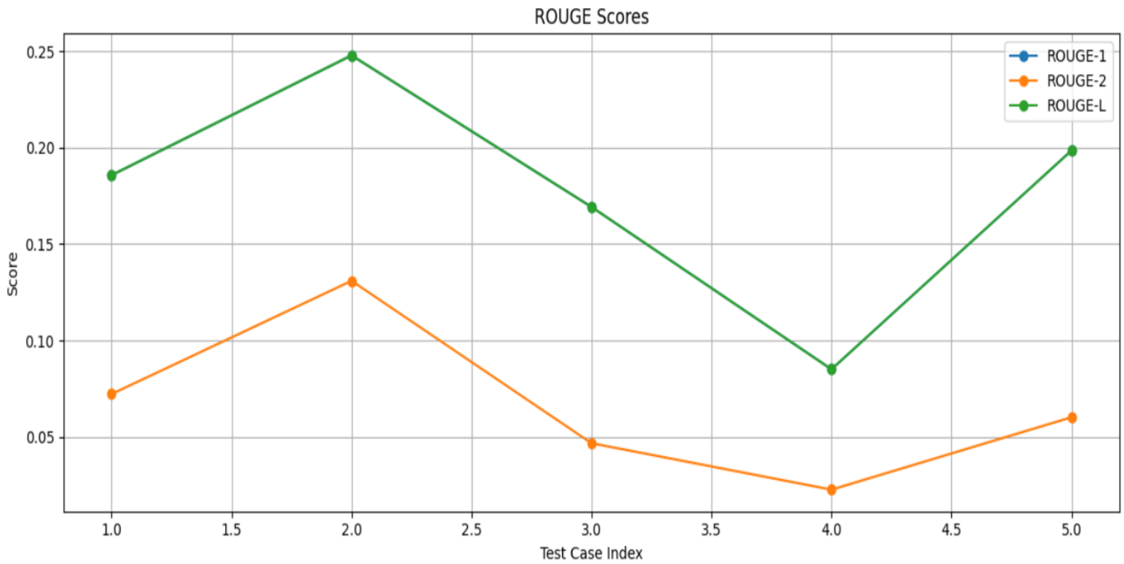

Figure 3.16: ROUGE Score Evaluation Results. This figure shows the performance of the chatbot using ROUGE metrics such as ROUGE 1, ROUGE 2, and ROUGE L.

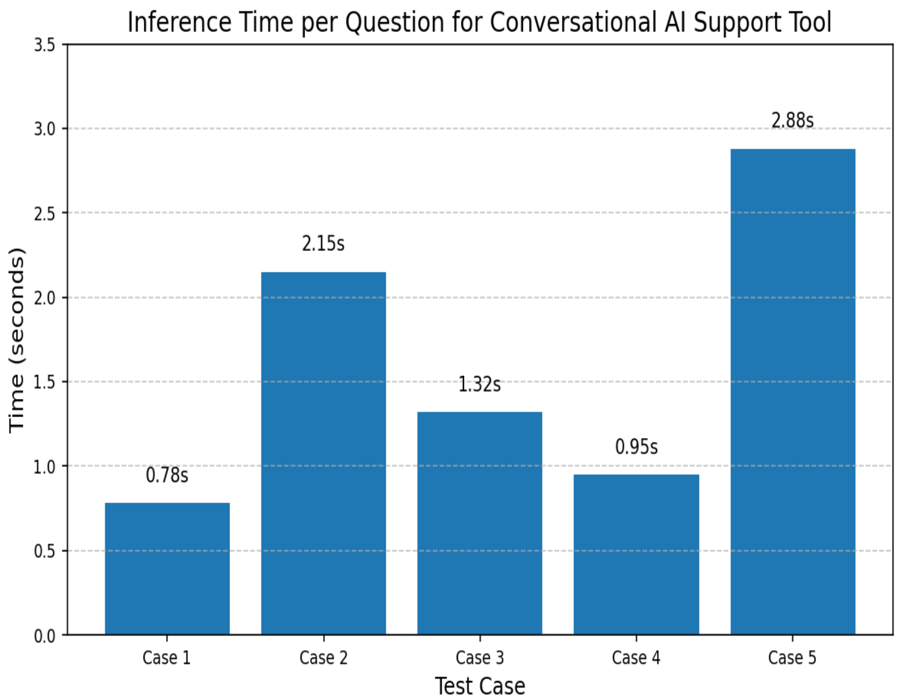

Figure 3.17: Average inference time (in seconds) per query. This graph illustrates the system's response latency, which directly affects user experience and real-time interaction quality. Lower values indicate better performance.

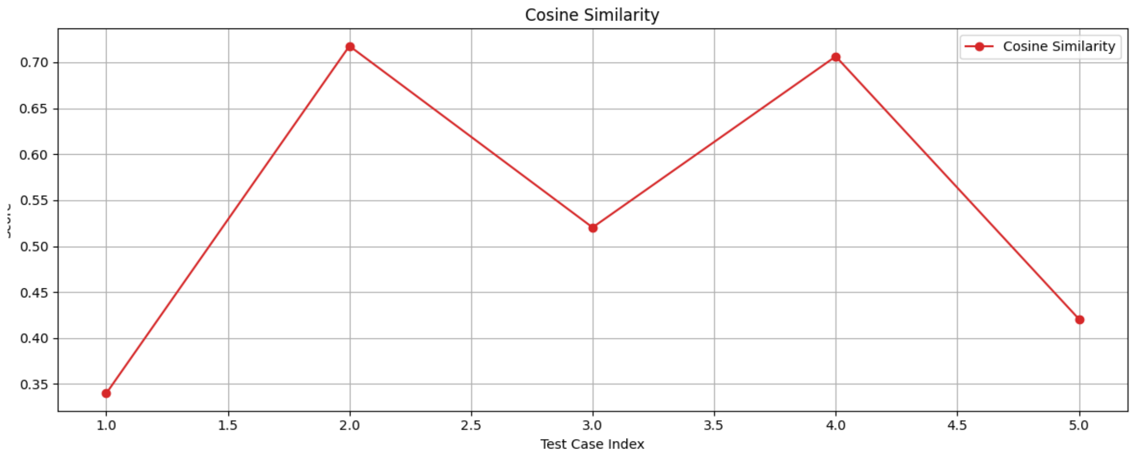

Figure 3.18: Cosine similarity scores of chatbot responses compared to ground truth responses. High similarity values suggest better semantic alignment and contextual relevance of the generated answers.

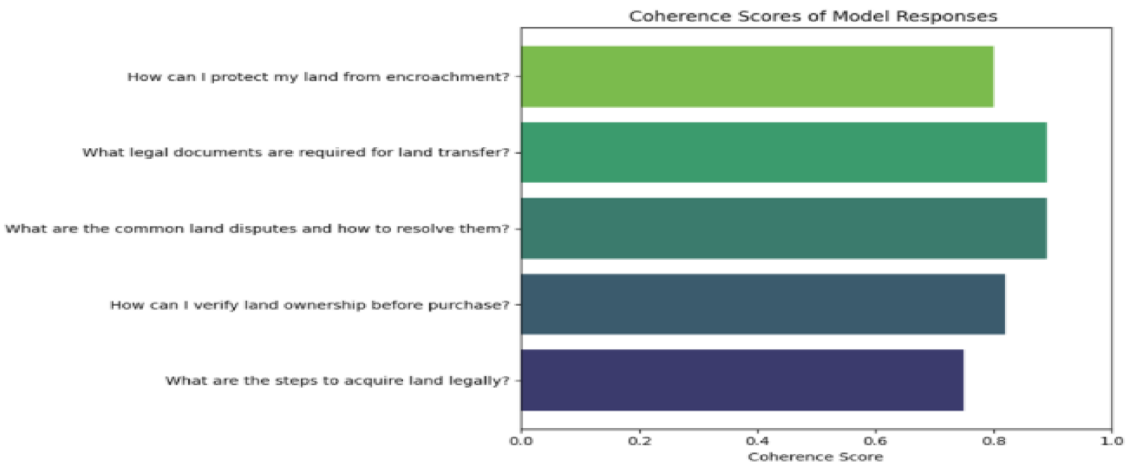

Figure 3.19: Coherence score evaluating the logical flow and fluency of the chatbot's responses. Higher coherence implies more natural and human-like language generation.

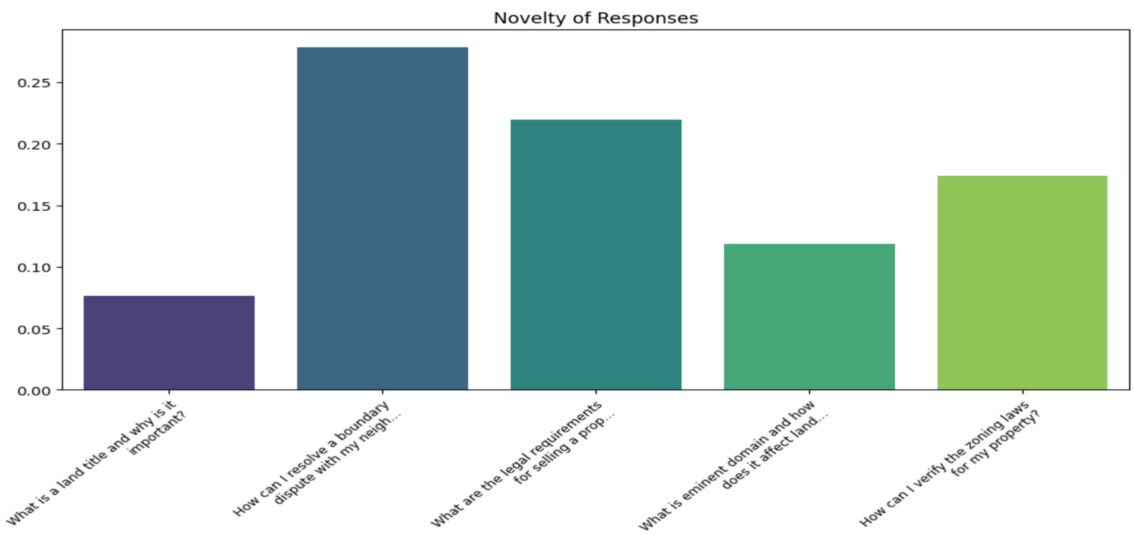

Figure 3.20: Novelty score of chatbot responses. This metric measures how much new, non-redundant information is provided in the responses. High novelty indicates creative and original answers, while maintaining relevance to the query.

3.6 Image Corpus Description

The image data component of this study focused on land ownership documents, specifically legal land titles commonly encountered in Uganda. These include various forms such as Freehold and Mailo titles, which are essential for determining tenure and property rights. Recognizing that end users are likely to capture and submit document images using their mobile phones, often under non-ideal conditions, the training process was designed to reflect these real-world scenarios.

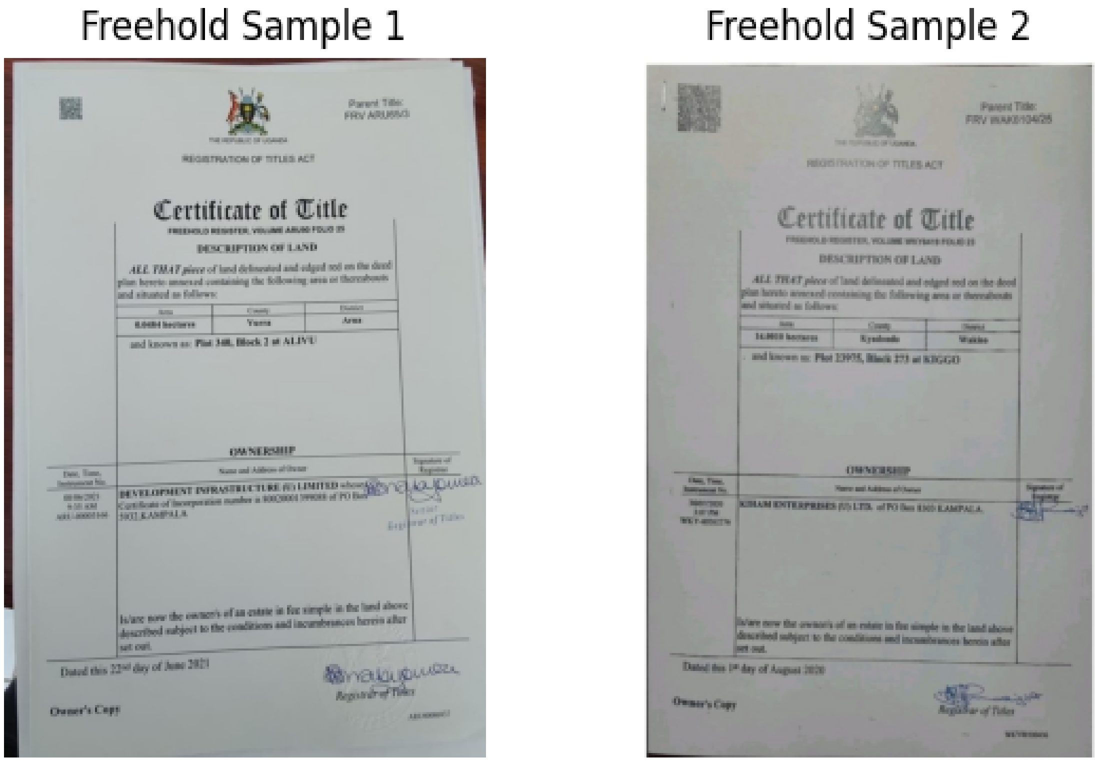

Figure 3.21: Samples of Freehold Certificates of Title.

Images used for model development were collected in formats that mimic typical user inputs, including variations in lighting, angles, resolution, and background clutter. To ensure that the computer vision model could generalize effectively across such diverse conditions, preprocessing techniques such as data augmentation were applied. These steps introduced variability and helped simulate common distortions and imperfections found in user-submitted photos. This realistic approach to data collection and preparation aimed to enhance the model's robustness and usability when deployed in practical legal advisory contexts.

The initial land title dataset comprised 61 images for Freehold Land Titles and 96 images for Mailo Land Titles. To address this class imbalance and prepare for effective model training, the dataset was balanced by oversampling the Freehold Land Titles class to match the Mailo Land Titles count, resulting in 96 images for each class and a balanced set of 192 unique images. To further improve the model's ability to generalize, a tenfold augmentation was applied to this balanced dataset. Augmentation techniques such as random rotations, flips, brightness adjustments, cropping, and noise injection were used to simulate diverse real-world conditions.

This process expanded the dataset to a balanced total of 1,920 images, with 960 images for each class. From this augmented pool, a final balanced dataset of 1,100 images was randomly selected, consisting of 550 images per class. This final dataset was then split into training (80%), testing (10%), and evaluation (10%) sets, ensuring robust and reliable model performance.

Figure 3.22: Samples of Mailo Certificates of Title

Freehold land titles as shown in Figure 3.21 grant the holders full ownership of the land in perpetuity, subject to the laws of Uganda. This form of tenure allows the owner to use, develop, lease, sell, or bequeath the land with minimal restrictions. Freehold tenure is considered one of the most secure land ownership systems, often issued following land registration under the Uganda Land Act. Mailo land titles as shown in Figure 3.22, on the other hand, are unique to Uganda and are rooted in the 1900 Buganda Agreement. This form of tenure grants perpetual ownership of land but typically coexists with the rights of lawful occupants, who may reside or cultivate the land. Mailo titles are prevalent in central Uganda and involve a more complex relationship between the registered owner and tenants, often requiring legal clarity during transactions or disputes. Both title types are legally recognized and are critical in establishing proof of ownership, facilitating transactions, and resolving land-related conflicts. For the system, understanding and distinguishing between these formats was essential in training computer vision models and preparing the image corpus to support automated land document analysis.

3.7 Image Corpus EDA

The image dataset used in this study consisted of a total of 161 annotated land title documents, categorized into two primary classes: 65 Freehold land titles and 96 Mailo land titles. This imbalance is shown in Figure 3.23. To address class imbalance, which can introduce bias during model training, oversampling was applied to the minority class, specifically the Freehold land titles, to ensure a more balanced representation within the dataset as seen in Figure 3.24.

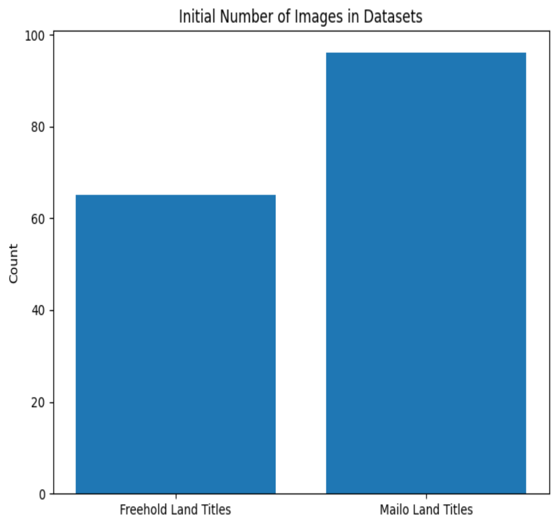

Figure 3.23: Image distribution in the original dataset showing class imbalance.

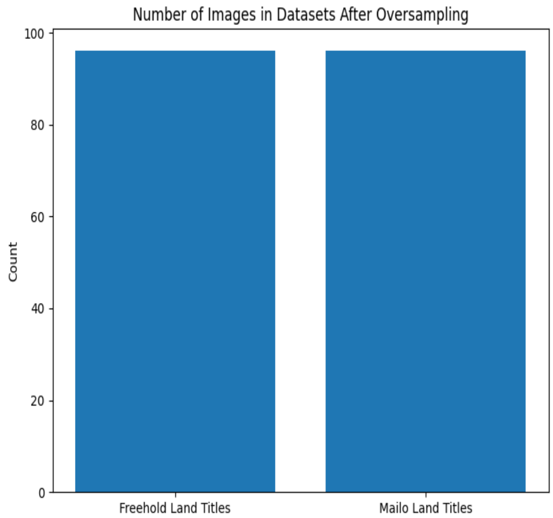

Figure 3.24: Image distribution after oversampling the minority class to achieve balance.

Pixel normalization was performed to ensure consistent intensity values across all images, reducing lighting variations and improving model convergence during training. The original pixel distribution shown in fig 3.25, displayed significant scatter, while Figure 3.26 shows a more concentrated distribution post-normalization. This standardization enabled the model to focus on meaningful visual features instead of inconsistencies in brightness or contrast.

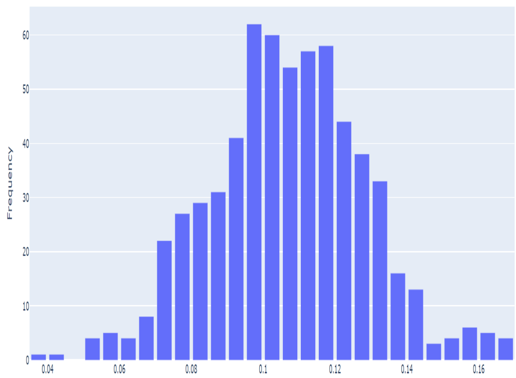

Figure 3.25: Pixel distribution before normalization. The values are unevenly spread across intensity levels.

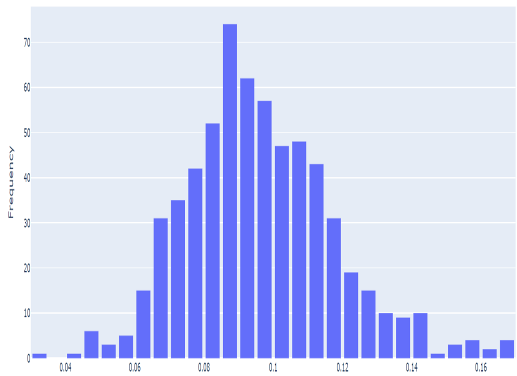

Figure 3.26: Pixel distribution after normalization. Intensity values are more standardized across the dataset.

The data augmentations (Figure 3.27 and Figure 3.28) were necessary to increase the diversity and volume of training samples, helping the model generalize better to real-world variations such as lighting, orientation, and image quality typical of phone captured land title photos. The Augmentation techniques included rotation, flipping and brightness adjustment.

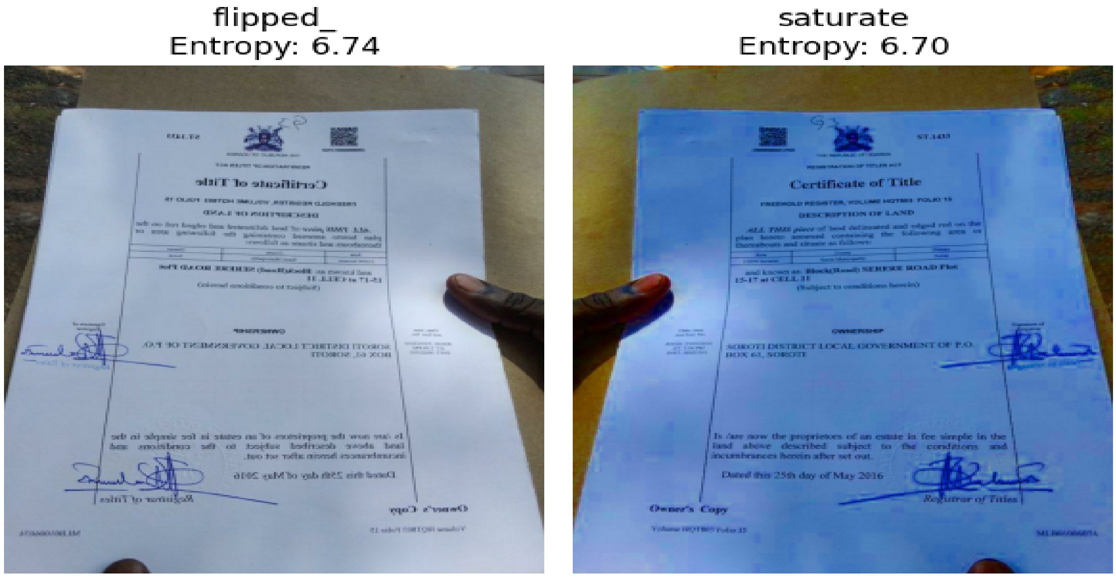

Figure 3.27: Image augmentations. flipped and saturated.

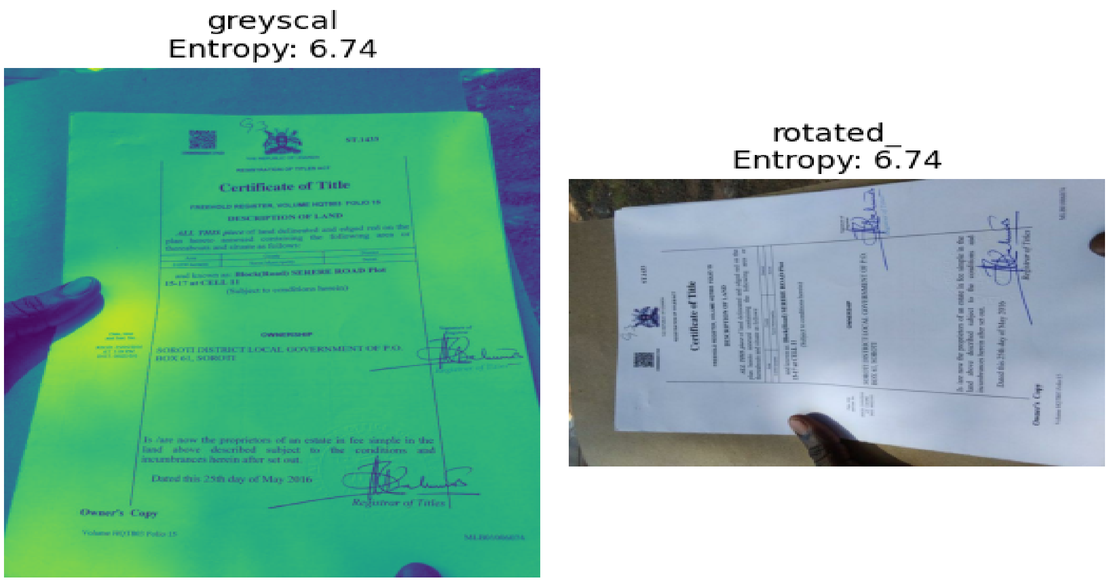

Figure 3.28: Image augmentations. greyscaled and rotated

The preprocessing pipeline was critical for enhancing the dataset's quality and reliability. Data augmentation not only addressed class imbalance but also improved the model's ability to generalize to diverse document layouts and scanning qualities. The final dataset, consisting of enhanced and balanced land title images, was used to train the computer vision model within the multimodal advisory system. Figure 3.29 showcases samples of genuine and counterfeit land title images used in training, aiding the model in distinguishing authentic from counterfeit documents effectively.

Figure 3.29 shows the corpus of real and fake images used to train the land title verification model. This corpus includes both original and counterfeit land certificates of title, enabling the model to learn critical visual distinctions. The ”Fake Land Titles” on the left and ”Original Land Titles” on the right represent the data fed into the image classification tool, which is designed to accurately differentiate between them.

Figure 3.29: Examples of original and fake land title images used in training to help the model learn to distinguish authentic from counterfeit land certificates of title.

3.8 Model Architecture

The visual architectures employed for processing scanned land titles in the legal land advisory system are shown below. The models are designed to extract and interpret features from the land titles to enable accurate and context-aware retrieval (Figure 3.30).

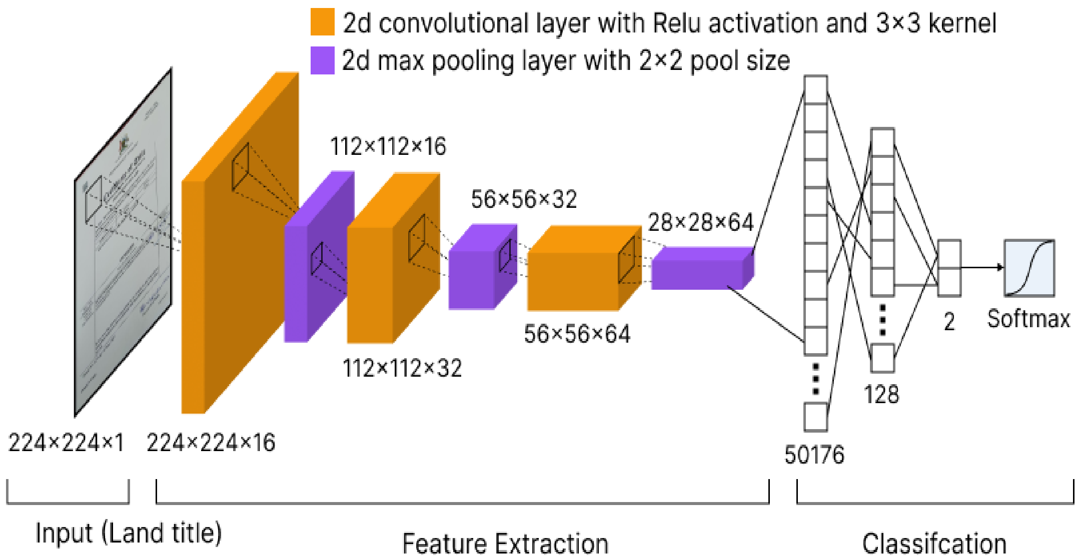

Figure 3.30: CNN architecture for extracting localized visual features from scanned land title documents, including seals, stamps, text blocks, and boundary diagrams.

As shown in Figure 3.30, the Convolutional Neural Network (CNN) architecture is employed to detect and encode spatially localized patterns in scanned land titles. This includes legal marks, handwritten text zones, and structural elements which are critical for document parsing in Uganda's land administration context.

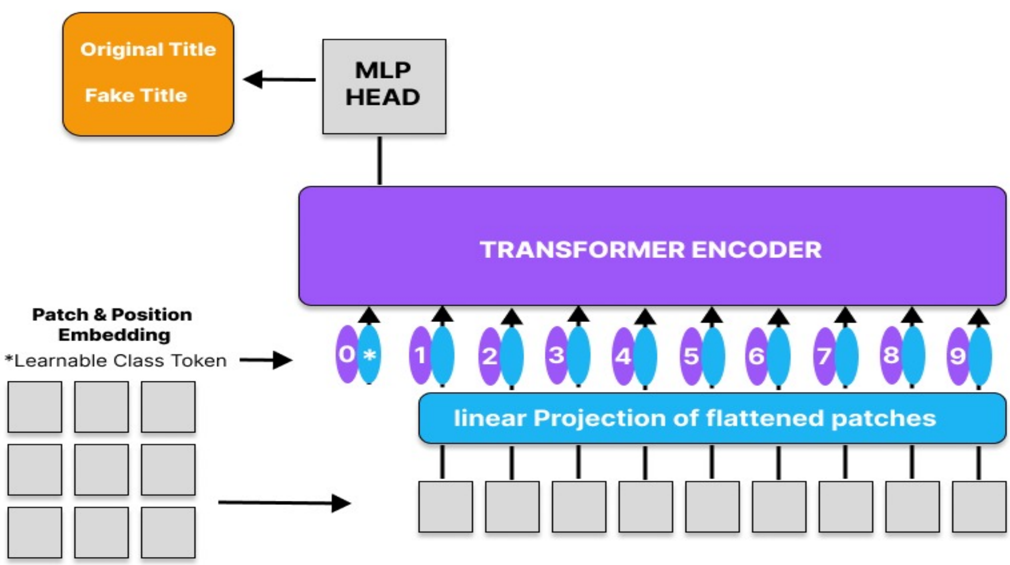

Figure 3.31: Vision Transformer (ViT) architecture for modeling long-range dependencies and document layout structures in land title images.

Figure 3.31 presents the Vision Transformer (ViT) architecture, which processes land title images as sequences of visual patches. This enables the model to reason over global layouts, such as boundary outlines, cadastral annotations, and spatial hierarchies commonly present in legal land documents.

3.9 Model Training

This section describes the training setup for the computer vision models used to classify land title images as either original or fake. Five models were trained on 50 epochs each. A custom Convolutional Neural Network (CNN), VGG16, VGG19, ResNet50, and Vision Transformer (ViT). Transfer learning was used for the pretrained architectures of VGG16, VGG19, ResNet50, and ViT. All models were trained using the balanced dataset of land titles, which was preprocessed and augmented to simulate real world mobile image capture conditions. The training process involved monitoring validation performance to prevent overfitting, and early stopping was applied where necessary. The training and validation curves for each model are shown in Figures 3.32, 3.33, 3.34, 3.35, and 3.36.

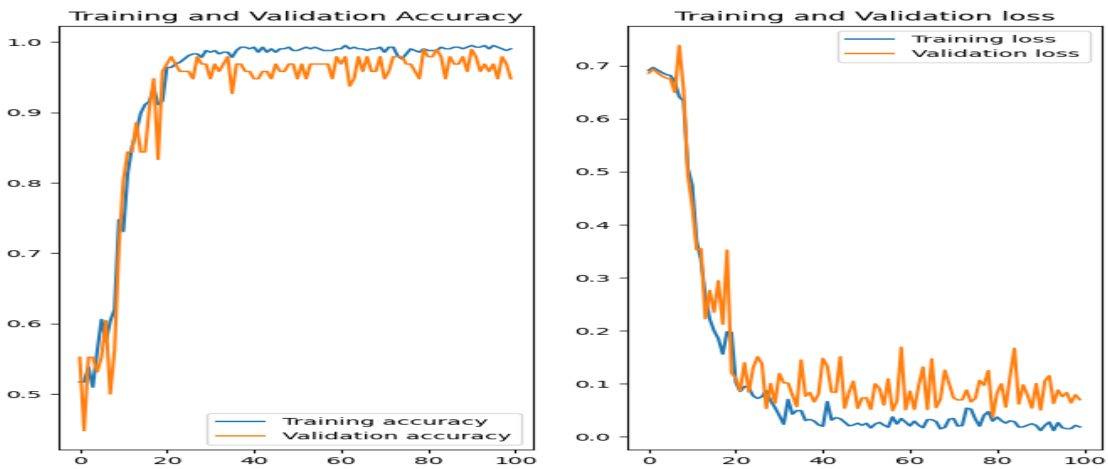

Figure 3.32: Training and validation performance of the CNN model, showing steady convergence across epochs.

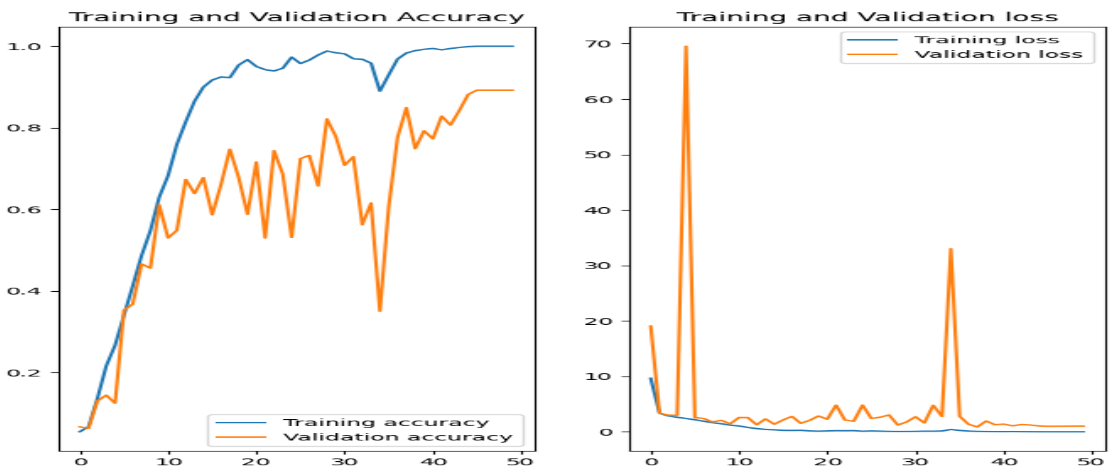

Figure 3.33: ResNet model training and validation curves, highlighting early improvements and generalization behavior.

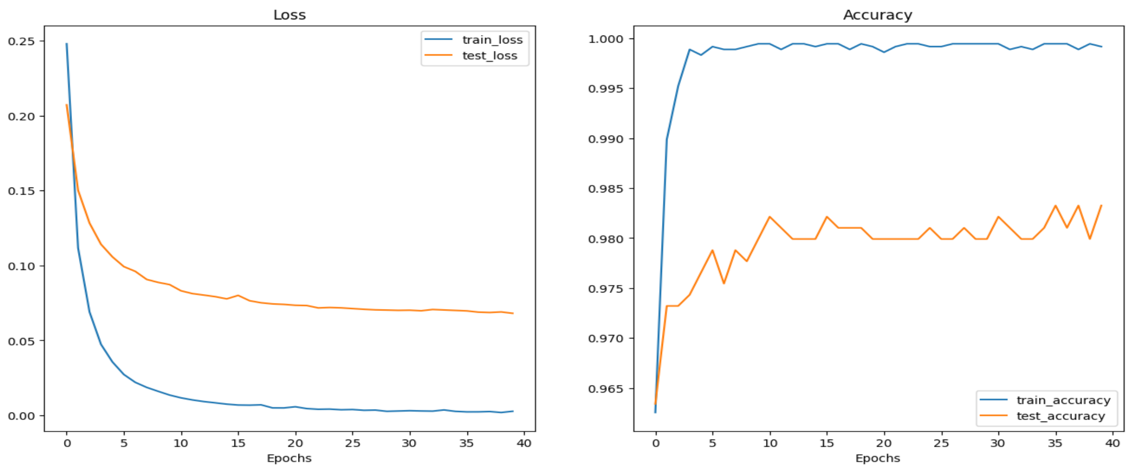

Figure 3.34: Vision Transformer (ViT) training and validation trends, indicating fluctuations in early training but strong final accuracy.

Figure 3.35: Training and validation metrics for the VGG16 model, showing progressive learning and reduced overfitting.

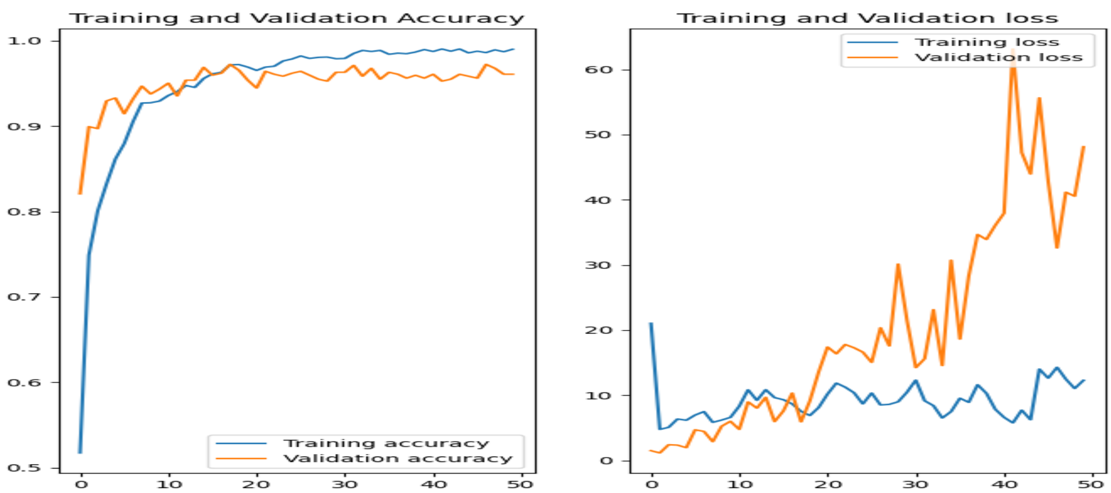

Figure 3.36: VGG19 model performance across epochs, with relatively stable training and validation trajectories.

3.10 Model Evaluation

Model evaluation was conducted to analyze the performance of each trained model. To assess the quality of predictions, standard classification metrics were computed. Accuracy, Precision, Recall, and F1 score. These metrics offer a view of the model's ability to correctly classify both original and fake land titles. Table 3.8 summarizes the final evaluation results of each model on the test dataset.

Model	Epochs	Training Loss	Training Accuracy
ResNet	50	0.15	0.95
CNN	50	0.01	0.99
VGG19	50	0.05	0.98
VGG16	50	1.50	0.97
ViT	50	0.02	0.99

Table 3.7: Training Phase Progress Metrics for Five Classification Models on the Land Title Dataset.

Figure 3.37: Training Set Performance Comparison showcasing Training Accuracy (Train Acc) and Validation Accuracy (Val Acc) for different classification models.

While the Vision Transformer (ViT) outperformed the other models in terms of raw accuracy and robustness, it required significantly longer training time and greater computational resources. The CNN model also demonstrated strong performance with high efficiency. Although the VGG models achieved competitive results, their large model size posed challenges for deployment and increased the associated costs.

Model testing was conducted to verify generalization performance and ensure robustness beyond the training phase. Table 3.8 shows the comparison between validation and test accuracy, highlighting how well each model generalized to unseen legal land queries.

Model	**Test Loss**	**Test Accuracy**	**Precision**	**Recall**	**F1 Score**
ResNet	0.87	0.87	0.88	0.86	0.86
CNN	1.01	0.97	0.97	0.96	0.96
VGG19	2.23	0.97	0.97	0.97	0.97
VGG16	0.92	0.89	0.89	0.88	0.88
ViT	0.45	0.98	0.98	0.98	0.98

Table 3.8: Test Set Evaluation Metrics for Five Classification Models on the Land Title Dataset.

The final evaluation shown in Figure 3.38 compared all models on the test dataset using key performance metrics to assess their overall performance. Figure 3.38 summarized each model's precision, recall, F1 score, and validation accuracy, providing a view of their strengths and weaknesses.

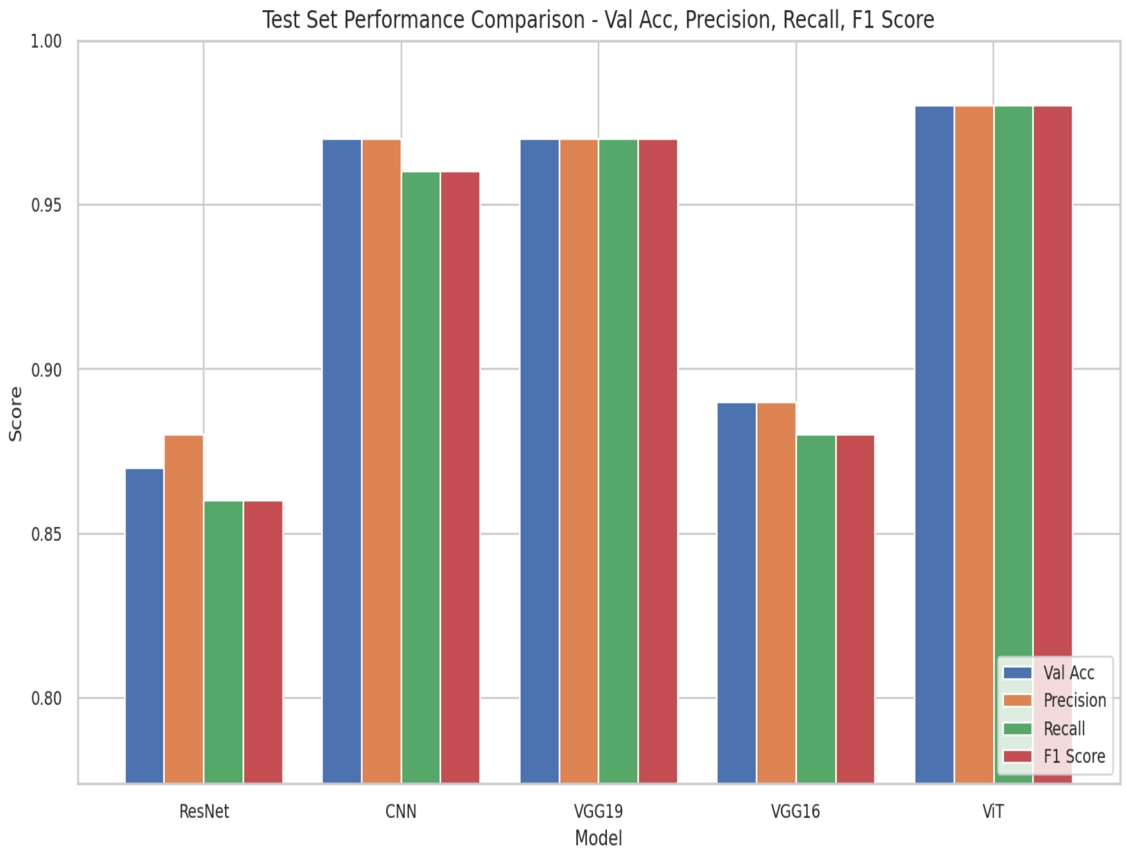

Figure 3.38: Test Set Performance Comparison of models across key metrics: Validation Accuracy (Val Acc), Precision, Recall, and F1 Score.

To validate our models' efficacy with real world applications, we proceeded to test their predictive accuracy against actual Ugandan land title certificates.

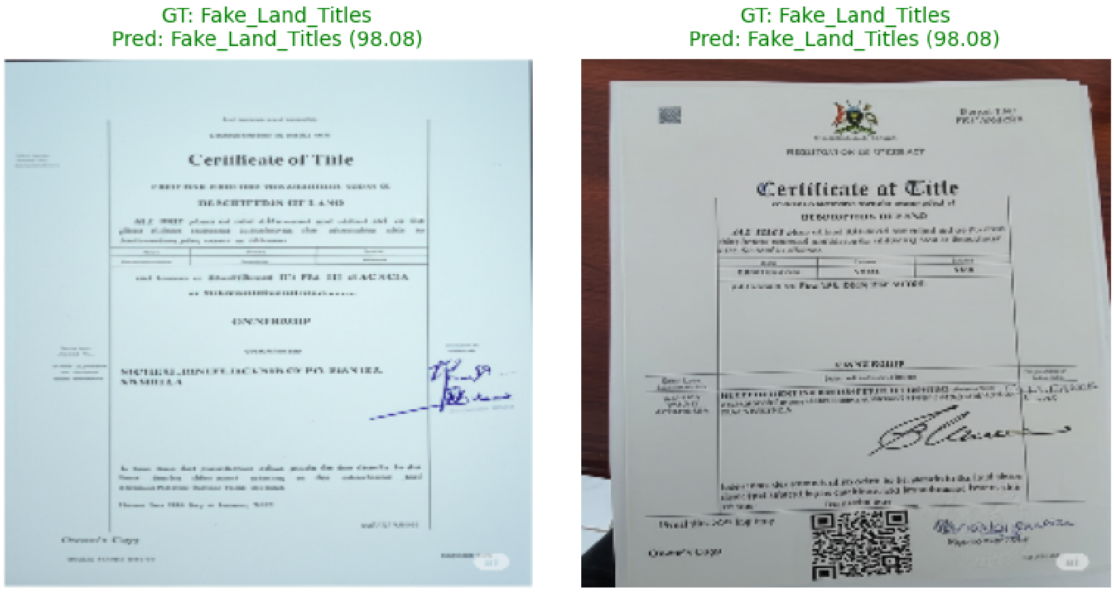

Figure 3.39: CNN model predictions on test images, showing effective classification of land title classes.

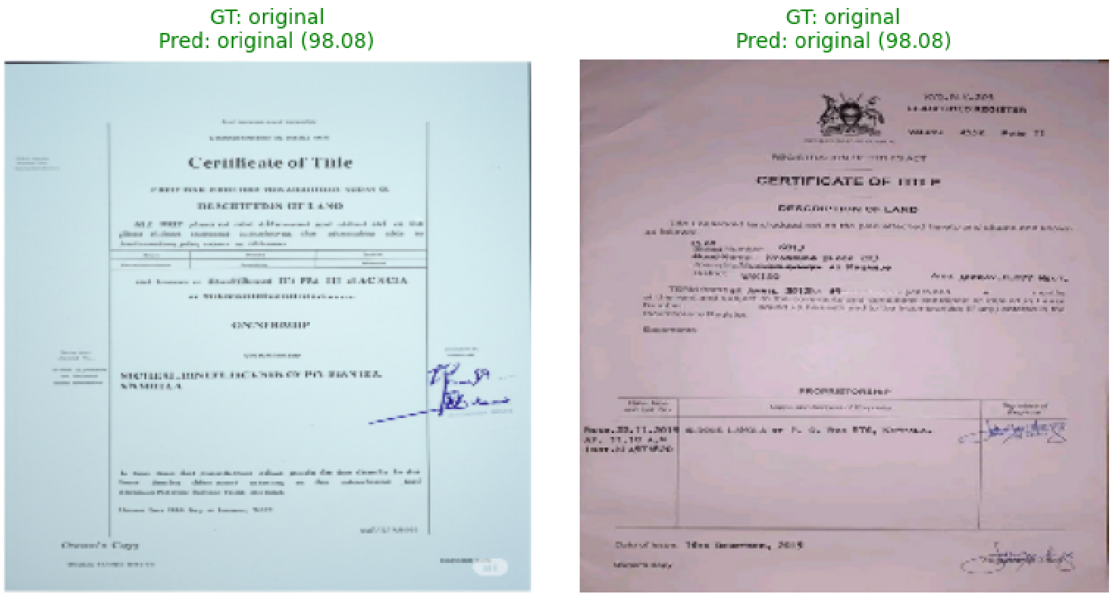

Figure 3.40: Vision Transformer (ViT) predictions on test images, highlighting its ability to capture fine grained features.

The confusion matrix provides a detailed breakdown of the model's classification performance by showing the number of correct and incorrect predictions for each class. High values along the diagonal indicate strong performance, where the predicted labels match the true labels. Off diagonal elements represent misclassifications.

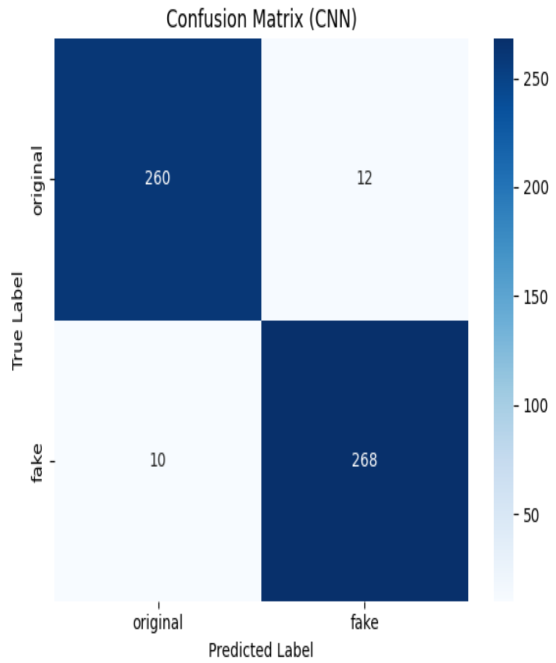

Figure 3.41: Confusion matrix of the CNN model showing strong classification performance across land title classes.

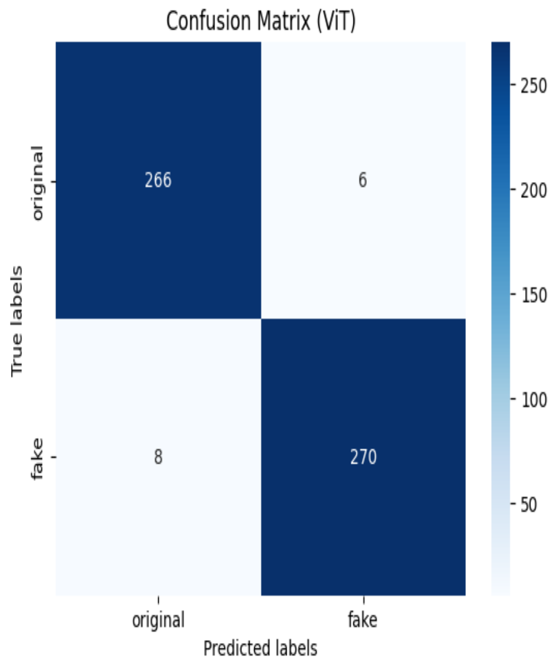

Figure 3.42: Confusion matrix of the ViT model indicating accurate and balanced predictions across classes.

In this study, the Vision Transformer (ViT) exhibited minimal confusion between classes, reflecting its superior ability to distinguish between land title categories. The matrix confirms that ViT achieved high accuracy and class level consistency, reinforcing its suitability for real world deployment in land use classification.

3.11 Language Translation

Spectrograms were used to analyze the generated speech, ensuring high fidelity and naturalness of responses, particularly with respect to the translated Luganda output. Spectrograms visually represent the frequency content of an audio signal over time, allowing us to assess phonetic clarity, pitch stability, and overall speech quality. This assessment is critical in validating the Text-to-Speech (TTS) component that converts generated legal text into spoken form, especially when integrated with the Luganda translation module.

To establish a baseline, a detailed spectrogram of the English TTS output was first generated. This visualization enabled focused analysis of the English audio's spectral features, including pronunciation patterns, energy distribution, and formant transitions ensuring that the synthesized voice was intelligible and consistent for legal advisory purposes.

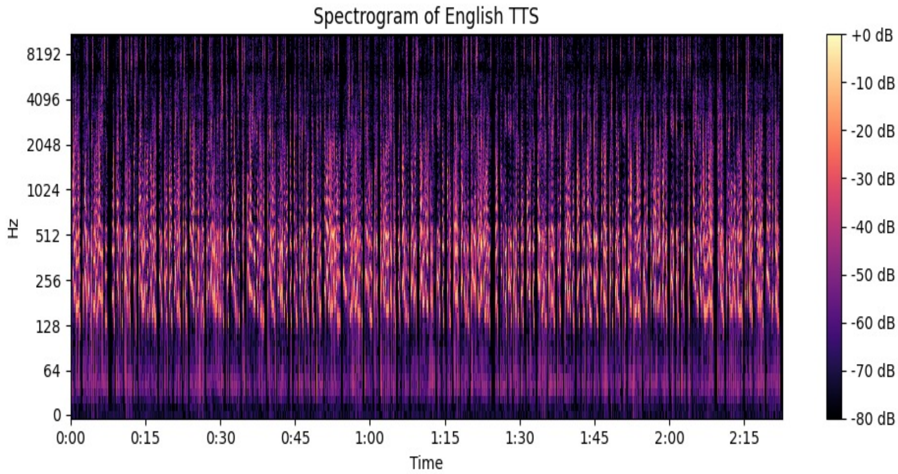

Figure 3.43: Spectrogram of Synthesized English TTS. This figure provides a detailed spectral representation of the English speech output, allowing for analysis of its acoustic features.

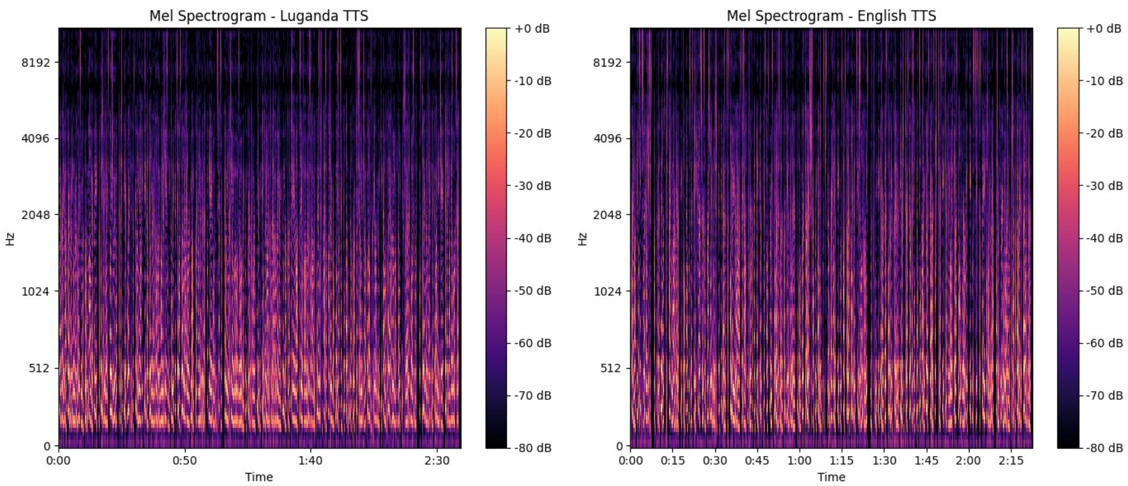

Figure 3.44: Comparative Mel Spectrograms of Synthesized Speech: Luganda TTS (left) vs. English TTS (right). This figure illustrates the frequency and intensity characteristics of speech generated in both languages, aiding in the assessment of synthesis quality.

Following this, Mel Spectrograms were generated for both Luganda and English synthesized speech (Figure 3.44). This comparative visualization provided insight into cross-lingual synthesis quality by highlighting variations in pitch contours, intonation flow, and spectral energy across both outputs. Given that Luganda is a tonal language, its accurate reproduction in TTS requires preserving nuanced pitch patterns that carry semantic weight. By examining both formant behavior and spectral consistency across languages, we identified key areas for model improvement, especially in rendering the tonal and prosodic features of Luganda speech. This process helped ensure that both English and Luganda audio outputs remain natural, clear, and suitable for delivering legally sensitive information in a conversational AI setting.

3.12 Explainable Model Predictions

To improve interpretability and transparency of the model's predictions, LIME (Local Interpretable Model Agnostic Explanations) was employed. LIME helps visualize which parts of the input most influenced the model's decision by highlighting regions or features that contributed positively or negatively to the predicted class. This is particularly useful for validating that the model is focusing on relevant land title content rather than spurious features.

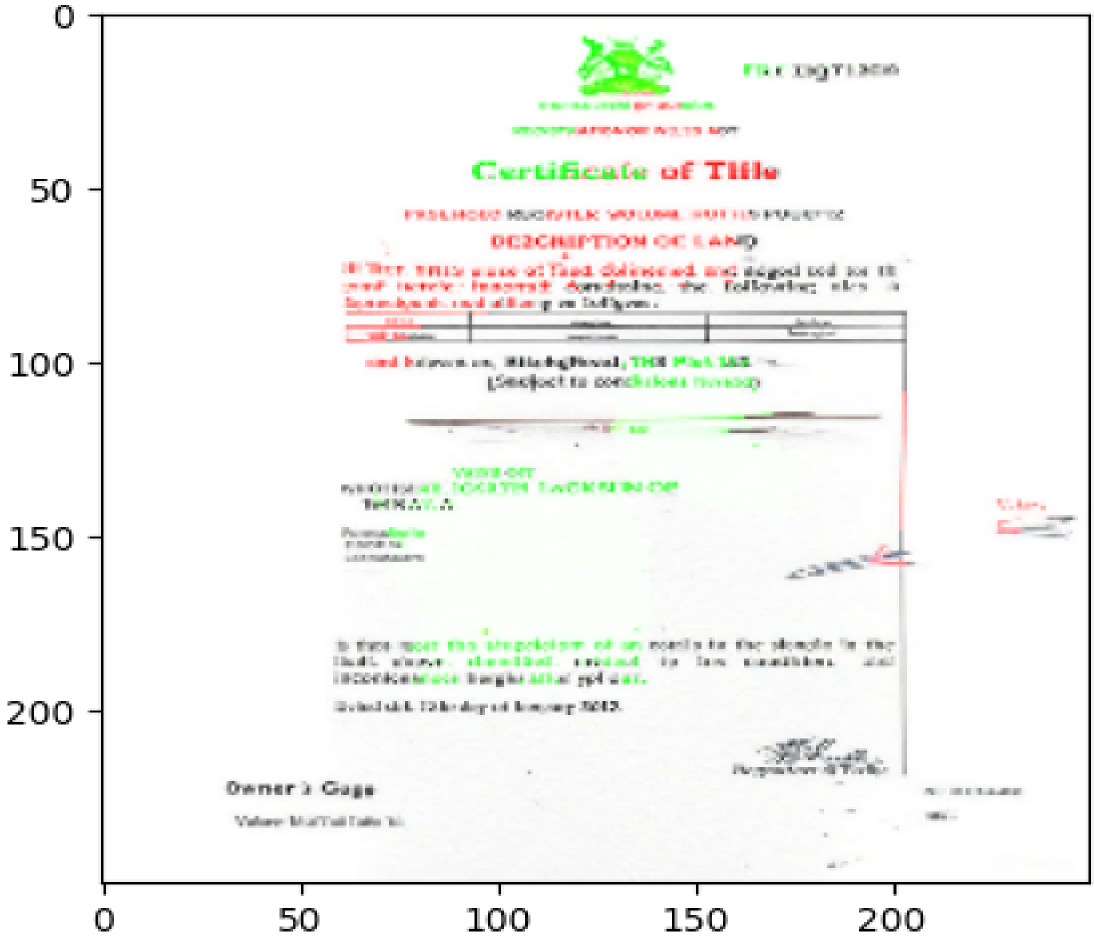

Figure 3.45: LIME explanation for the CNN model highlighting important image regions influencing predictions.

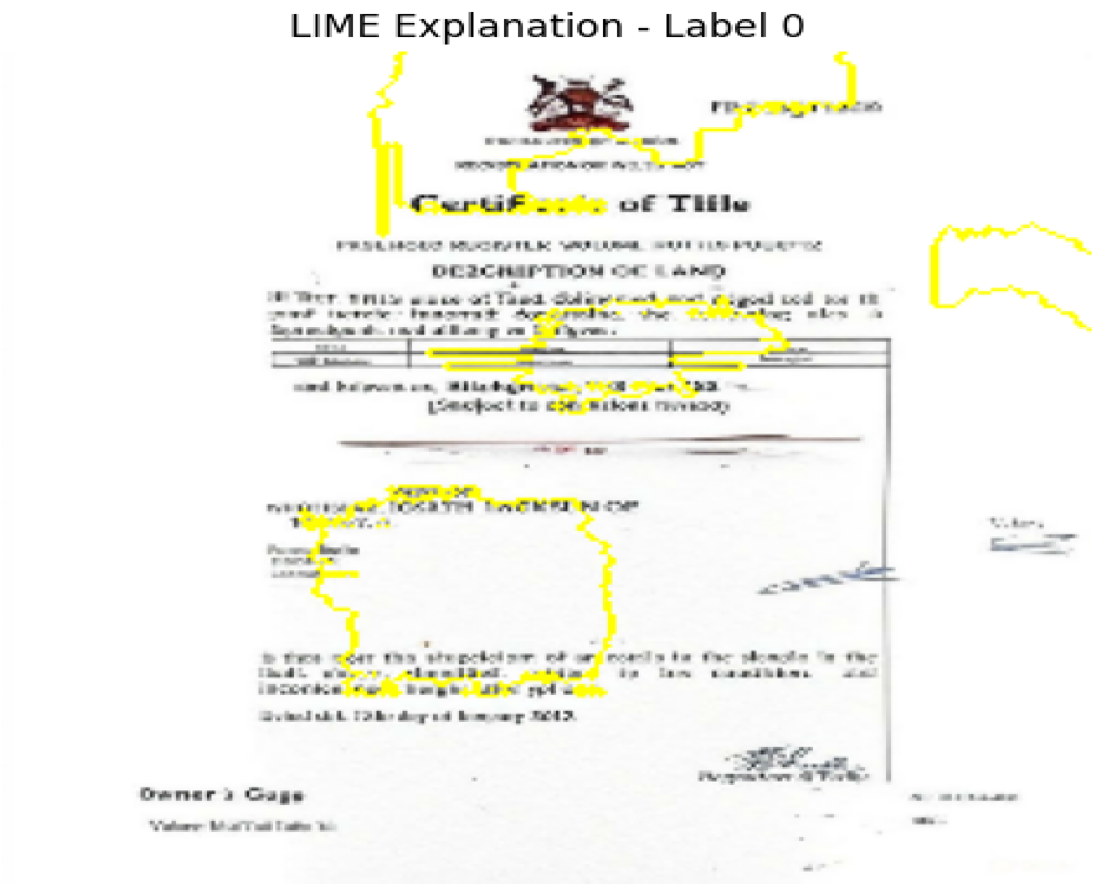

Figure 3.46: LIME explanation for the CNN model showing perturbations that highlight key image regions influencing predictions.

As shown in the visualizations, LIME successfully highlights critical textual or visual elements that align with human expectations, offering insights into the decision making process of complex models such as CNN, VGG, and ViT. SHAP (SHapley Additive exPlanations) was applied to the best performing models to generate detailed visual explanations for individual predictions. This helped validate that the models focused on relevant image areas, such as key text and security features on land titles, increasing trust in the model's outputs. The SHAP visualizations complemented other explainability methods like LIME and Grad CAM, offering both global and local insights into the decision-making process.

Figure 3.47: SHAP explanation for the CNN model illustrating the contribution of different image regions to the model's prediction. Warmer colors indicate areas with a stronger positive impact on the output.

Model Selection And MLOPs.

4.1 Model Selection.

From the results shown in Table 3.8, the Vision Transformer (ViT) and CNN models were the best performers. The ViT model worked well because of its self-attention, which helps it understand the whole image. The CNN model also gave good results and was faster to train with less computing power. To maximize the strengths of both CNN and ViT models, we adopt an ensemble architecture illustrated in Figure 4.1.

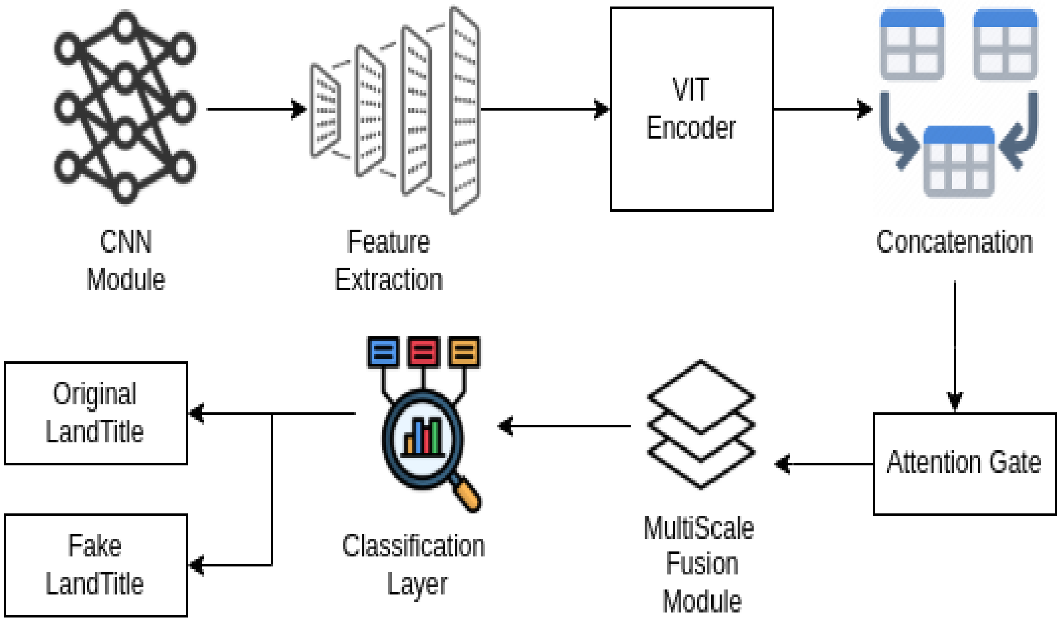

Figure 4.1: Ensemble architecture combining CNN and ViT encoders for robust land title analysis, integrating local and global visual features.

The CNN module captures fine-grained visual cues like document texture and signature blocks, while the ViT module models spatial relationships across the entire document. The fused visual representation is passed to the retrieval module within the RAG framework, enhancing semantic alignment between user queries and complex legal land documents.

4.2 Machine Learning Operations (MLOPs)

Machine Learning Operations (MLOps) ensured the scalability, reliability, and maintainability of the multimodal legal advisory system across its development and deployment lifecycle. Version control using Git was implemented to manage source code, track changes, and facilitate collaboration. Experiment tracking was handled through structured logging of training parameters and performance metrics, enabling reproducibility and comparative evaluation of different model versions. Data and model artifacts were versioned to ensure traceability, particularly for the translation and computer vision components. Models were tested using pre-defined datasets before deployment to reduce inference errors. Deployment was conducted using Streamlit for local development and testing, and Hugging Face Spaces for public hosting and access. These platforms allowed rapid prototyping while reducing infrastructure overhead. The deployment environment utilized Docker containers to ensure consistency and environment reproducibility. The modular architecture of the system allowed independent updates of the retrieval, translation, and vision components, aligning with MLOps best practices. Scalability was considered through the use of lightweight models and modular APIs, making it easier to adapt for offline or edge-device deployment. Overall, MLOps principles supported the entire model lifecycle, from training to versioned deployment, ensuring that the system remains robust, reproducible, and adaptable for future extensions.

Results

5.1 Results from the Deployed Chatbot

The captured interactions with the deployed Legal Land Advisory tool were captured in this results section. The purpose is to illustrate the system's capability to process and respond to user inquiries relevant to land acquisition, legal advisory services, and related matters within the context of Uganda.

Three sample queries were submitted to the chatbot to demonstrate its performance. The first query requested information on the procedural steps involved in acquiring land in Uganda. The second query inquired about the definition and scope of legal land advisory services. The third query asked about the cost of land in Uganda, where the response included a Luganda audio output alongside the English text. The system's responses to these queries are displayed in Figures 5.1, 5.7, and 5.3, respectively.

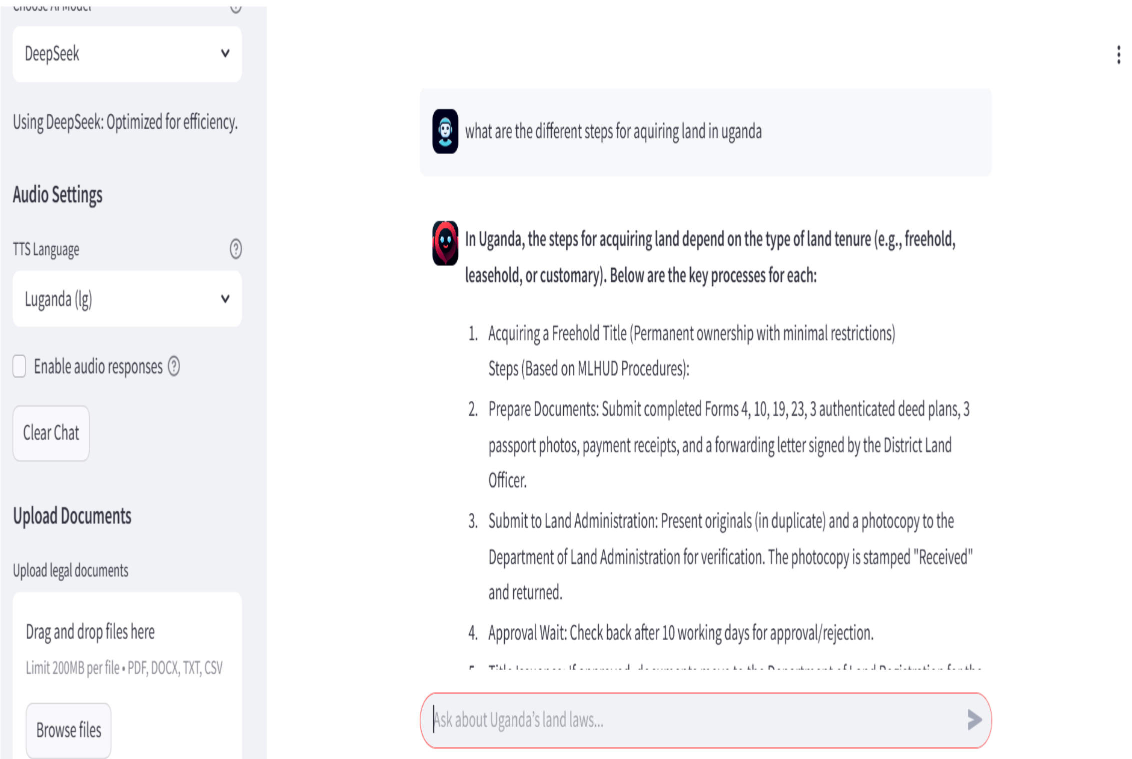

Figure 5.1: Chatbot response to the query regarding the steps for acquiring land in Uganda.

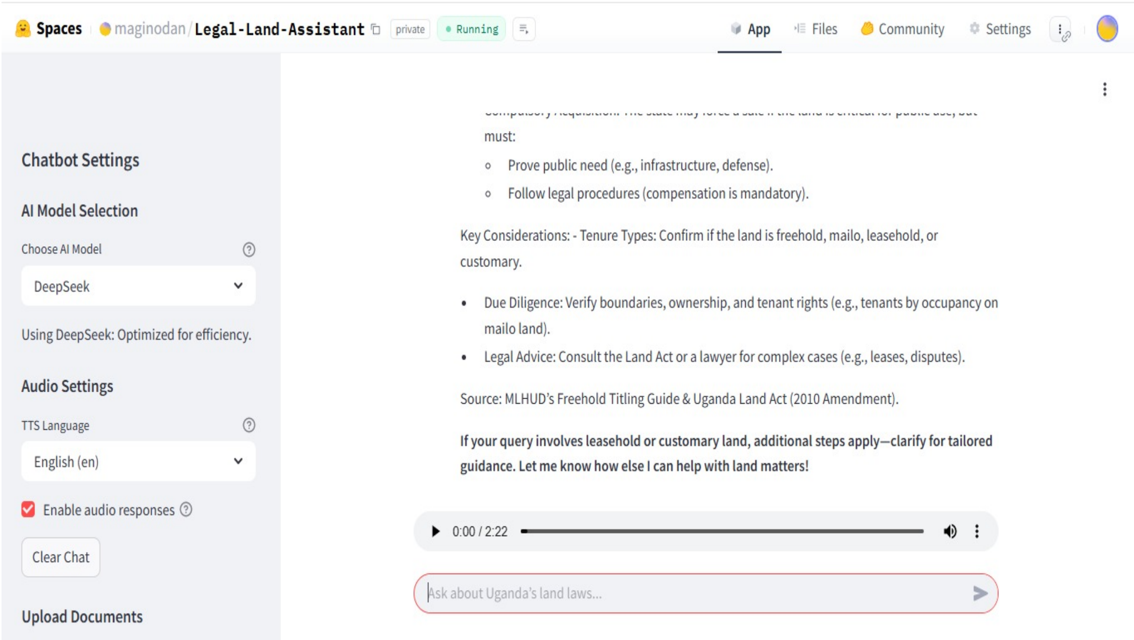

Figure 5.2: streamlit UI deployed on Hugging Face.

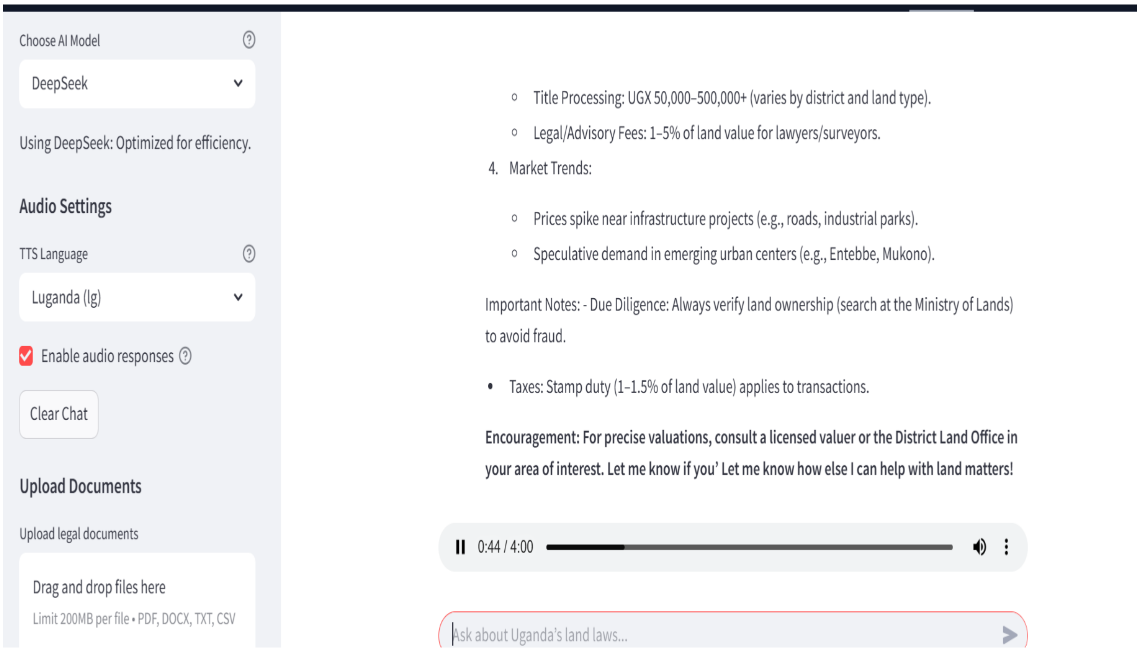

Figure 5.3: Chatbot response to the query about land cost in Uganda, with English text and Luganda audio support.

These screenshots provide direct evidence of the chatbot's ability to interpret natural language input and generate appropriate, contextually relevant responses upon deployment. The inclusion of these interaction examples serves to confirm the operational functionality of the chatbot in a real-world environment. The results further indicate that the system successfully integrates domain-specific knowledge, multilingual support, and user interface components to facilitate effective legal advisory communication.

5.2 Land Title Verification Tool.

The tool was able to correctly classify the authenticity of scanned land title documents by analyzing visual features such as seals, stamps, and text blocks. Furthermore, its explainable AI (XAI) capabilities, specifically utilizing Grad-CAM, provided visual explanations by highlighting regions within the document images that were most influential in the model's classification decision.

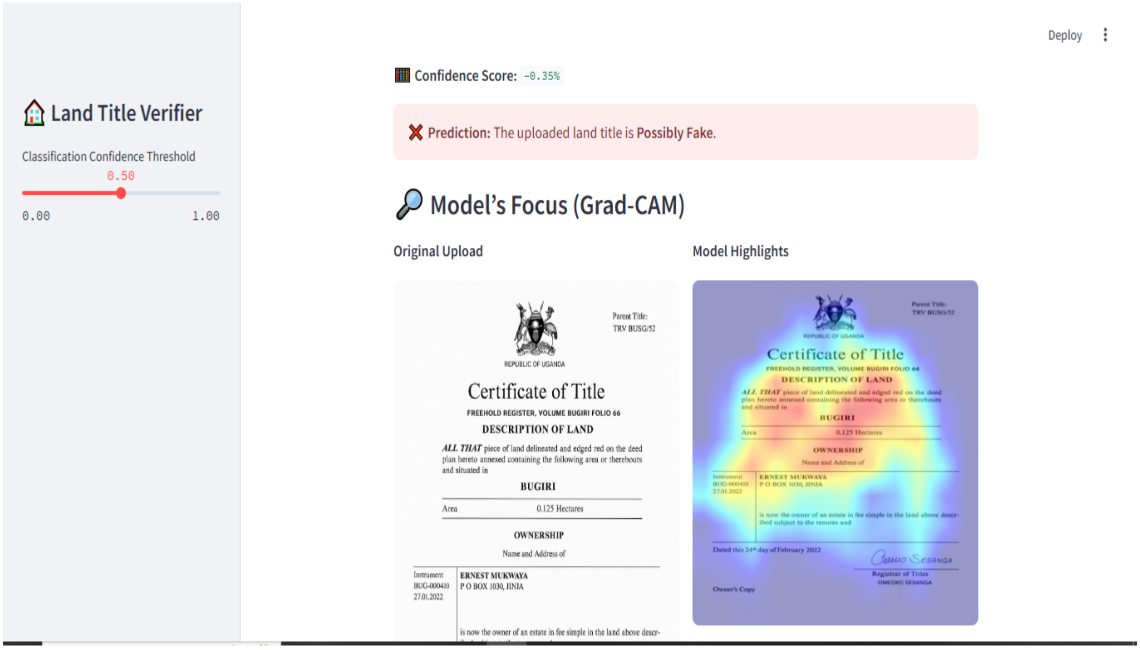

Figure 5.4: A fake land title was correctly classified as Forged

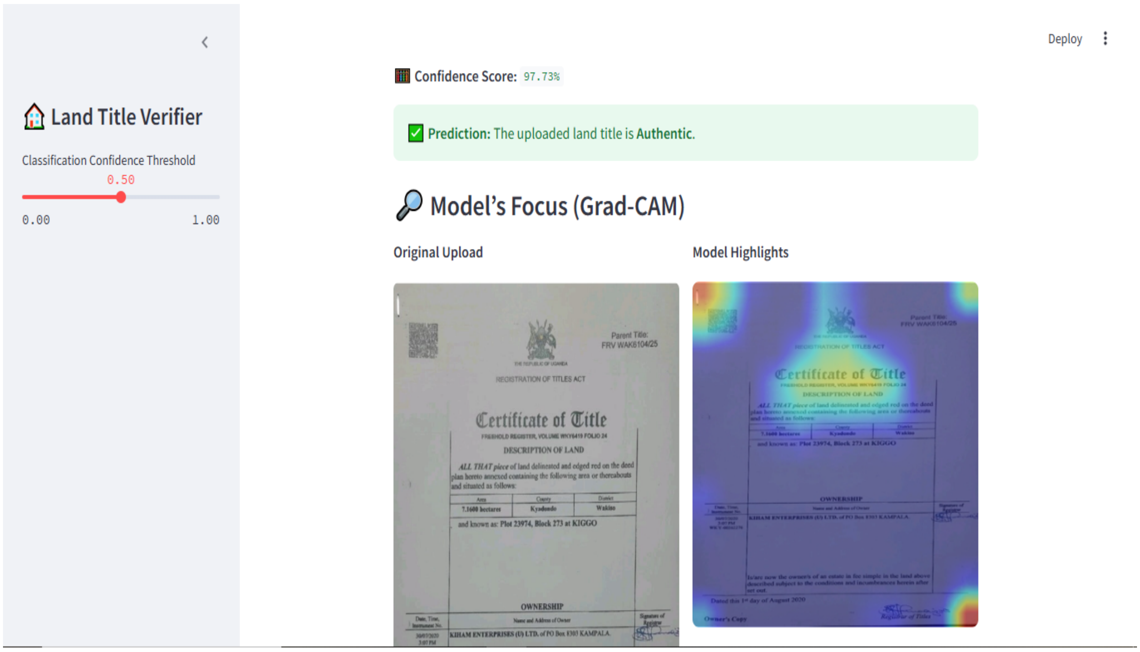

Figure 5.5: A genuine land title was correctly classified as genuine

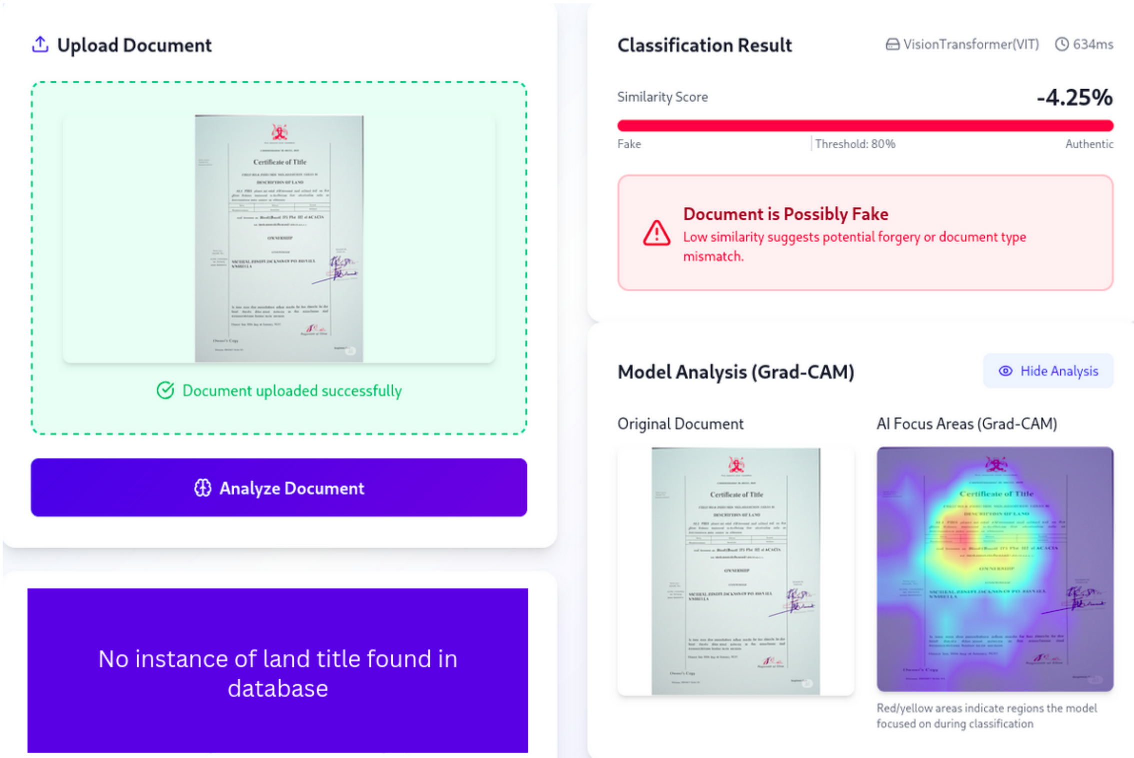

Figure 5.6: A land title was classified as fake and could not be located in any Land database as well further verifying the predictions of the model.

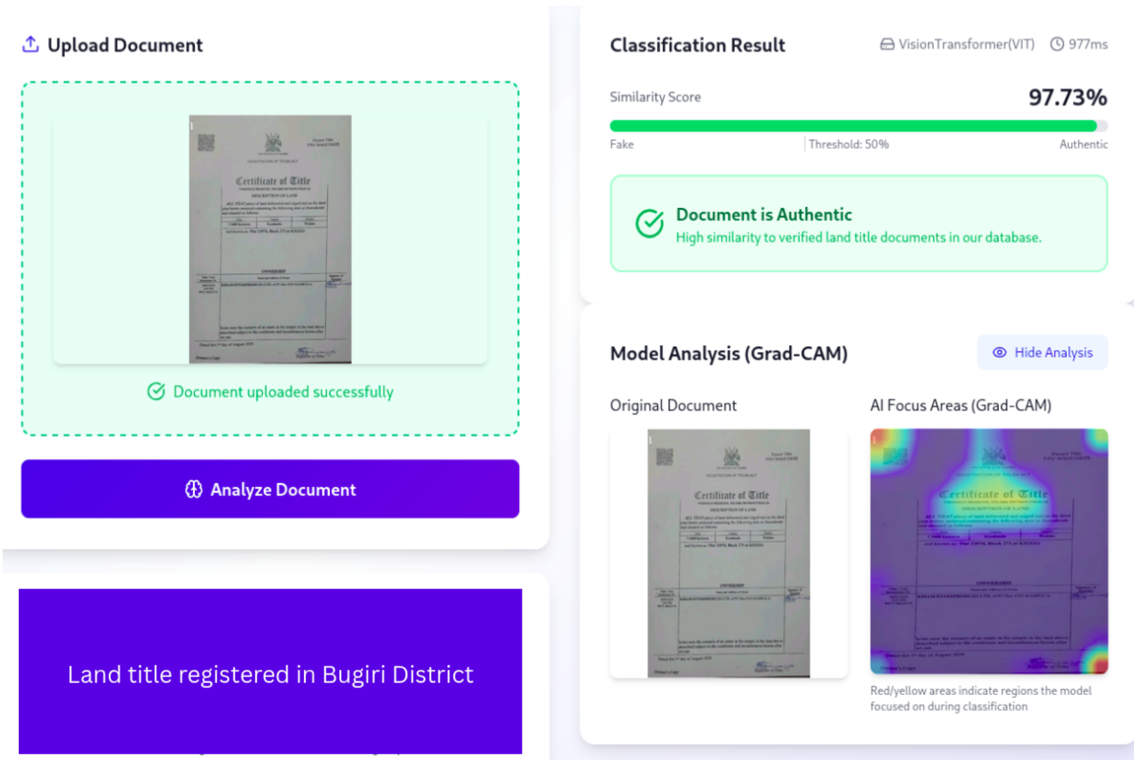

Figure 5.7: A Land title was classified as original and was found in the database as well further justifying the predictions of the model.

Limitations.

6.1 Limitations

Although data augmentation increased variety, the lack of a diverse and extensive real-world dataset hinders the model's ability to generalize to unseen document formats, especially handwritten or poorly scanned land forms. The multilingual translation system lacks a comprehensive parallel corpus for local languages such as Luganda and Swahili in the legal domain. This affects translation quality and the handling of contextually sensitive legal terms, thereby reducing trust in the accuracy of the output. Hosting the system on Hugging Face provides accessibility but imposes technical constraints, including memory usage limits, response latency, and dependence on a consistent internet connection. These factors impact the user experience, particularly in low-resource or offline settings. The system also relies on third-party APIs for embeddings, vector retrieval, and chatbot generation. This dependency introduces limitations for large-scale deployment, increases operational costs, and restricts the ability to operate in secure or offline environments.

Conclusions And Future Works.

7.1 Conclusions and Future Work

This study developed a multi-modal Conversational AI system to provide accessible legal land advisory services by combining Graph Retrieval Augmented Generation (GRAG), computer vision, and multilingual capabilities. By integrating domain specific legal documents into an interactive chatbot, the system enables natural language querying of complex legal content. Key innovations include exploratory data analysis, document parsing, translation, image classification, and conversational deployment, all of which aim to empower under served communities to navigate land rights and procedures. Local language support and visual document understanding enhance the usability of the system for people with limited literacy or technical experience.
Future work will focus on expanding annotated image datasets to improve computer vision model accuracy. To address translation challenges, a curated parallel corpus of legal texts in Luganda, Swahili, and English will be developed, enhancing neural translation models. Additionally, transitioning to hybrid or on-device deployment models will ensure greater accessibility and resilience in offline or resource constrained settings.

References

[1] C.-Y. Lin and F. J. Och, “Automatic evaluation of machine translation quality using longest common subsequence and skip-bigram statistics,” *Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics (ACL 2004)*, 2004.

[2] Legal Services Board, *Legal services board static market analysis*, Available at: https://research.legalservicesboard.org.uk/analysis/supply/sta, 2014.

[3] “Squad: 100,000+ questions for machine comprehension of text,” in *Proceedings of EMNLP*, 2018, pp. 2383–2392.

[4] N. Craswell, “Mean reciprocal rank,” *Encyclopedia of Database Systems*, L. Liu and M. T. Özsu, Eds., pp. 1703–1703, 2019.

[5] “An overview of chatbot technology,” in *Artificial Intelligence Applications and Innovations*, vol. 584, 2020, pp. 373–383.

[6] “Retrieval-augmented generation for knowledge-intensive nlp tasks,” in *Advances in Neural Information Processing Systems*, vol. 33, 2020, pp. 9459–9474.

[7] “Conversational ai applications in urban planning and land management,” *Journal of Urban Technology*, vol. 28, no. 4, pp. 95–115, 2021.

[8] “Improving legal dialogue systems with reinforcement learning,” *Computational Intelligence and Neuroscience*, vol. 2021, pp. 1–12, 2021.

[9] M. Ballout, U. Krumnack, G. Heidemann, and K.-U. Kuhnberger, “Investigating pre-trained language models on cross-domain datasets, a step closer to general ai,” *2023*, 2023.

[10] C.-Y. Lin and E. H. Hovy, “Automatic evaluation of summaries using n-gram co-occurrence statistics,” *Proceedings of the 2003 Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics (HLT-NAACL 2003)*, 2023.

[11] Private Sector Foundation Uganda (PSCP II), *Final draft issues paper – registration of titles act*, Jan. 2023.

[12] M. Breitkopf, “Augmenting large language model’s knowledge using external sources without finetuning,” Ph.D. dissertation, Ruhr-Universität Bochum, 2024.

[13] J. Chen, H. Lin, X. Han, and L. Sun, “Benchmarking large language models in retrieval-augmented generation,” in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 38, 2024, pp. 17 754–17 762.

[14] Z. Chen, “Efficient retrieval-augmented generation,” Ph.D. dissertation, University of Illinois at Urbana-Champaign, 2024.

[15] Daily Monitor, “African judges should prioritise case backlog,” Sep. 2024, Accessed: January 14, 2025. [Online]. Available: https://www.monitor.co.ug.

[16] Daily Monitor, “Commercial court case backlog locks up shs85t,” Feb. 2024, Accessed: January 14, 2025. [Online]. Available: https://www.monitor.co.ug.

[17] Daily Monitor, “Land wrangles top major legal problems, says report,” Mar. 2024, Accessed: January 14, 2025. [Online]. Available: https://www.monitor.co.ug.

[18] K. Guan, Q. Cao, Y. Sun, X. Wang, and R. Song, “Bsharedrag: Backbone shared retrieval-augmented generation for the e-commerce domain,” *arXiv preprint arXiv:2409.20075*, 2024.

[19] B. Han, T. Susnjak, and A. Mathrani, “Automating systematic literature reviews with retrieval-augmented generation: A comprehensive overview,” *Applied Sciences*, vol. 14, no. 19, p. 9103, 2024.

[20] Z. Jiang, X. Ma, and W. Chen, “Longrag: Enhancing retrieval-augmented generation with long-context llms,” *arXiv preprint arXiv:2406.15319*, 2024.

[21] S. Kukreja, T. Kumar, V. Bharate, R. Singh, A. Dasgupta, and D. Guha, “Impact of ensemble of vector embeddings on speculative retrieval augmented generation,” in *2024 5th International Conference on Data Intelligence and Cognitive Informatics (ICDICI)*, 2024, pp. 963–969. DOI: 10.1109/ICDICI62993.2024.10810945.

[22] B. Peng, Y. Zhu, Y. Liu, *et al.*, “Graph retrieval-augmented generation: A survey,” *arXiv preprint arXiv:2408.08921*, 2024.

[23] E. Sun, Y. Wang, and L. Tian, “Block-attention for efficient rag,” *arXiv preprint arXiv:2409.15355*, 2024.

[24] S. Wu, Y. Xiong, Y. Cui, *et al.*, “Retrieval-augmented generation for natural language processing: A survey,” *arXiv preprint arXiv:2407.13193*, 2024.

[25] S.-Q. Yan, J.-C. Gu, Y. Zhu, and Z.-H. Ling, “Corrective retrieval augmented generation,” *arXiv preprint arXiv:2401.15884*, 2024.

[26] S. Kumar, N. Sharma, R. Kumar, and S. Singh, “Gernerative ai & llms,”

Appendices

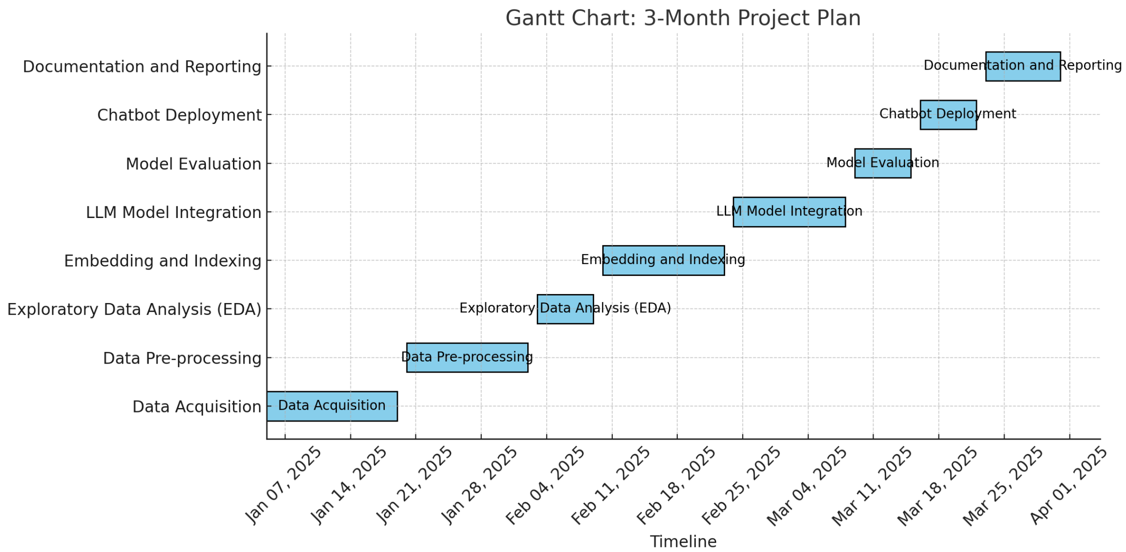

Figure 7.1: Gantt Chart showing Project completion deadlines

ERD

Budget Item	**Description**	**Monthly Cost (UGX)**	**Total Cost (UGX)**
Data Acquisition	Initial purchase of legal document access, plus maintenance of sources	75,000	600,000
Hosting and Cloud Services	Monthly hosting, data storage on AWS or Google Cloud, and server costs	60,000	480,000
API Integration	API usage costs e.g. OpenAI API subscription, Gemini, Hugging Face	60,000	480,000
UI/UX Design	Initial design phase with incremental improvements over the project	10,000	800,000
Dissemination and Reporting	Midterm and final reports, plus minor outreach events throughout the project	10,000	800,000
Publication	Fees for publishing research and findings	–	1,110,000
Ethical Clearance	Fees for obtaining ethical clearance	–	180,000
Total		**215,000**	**3,000,000**

Table 7.1: Detailed Budget

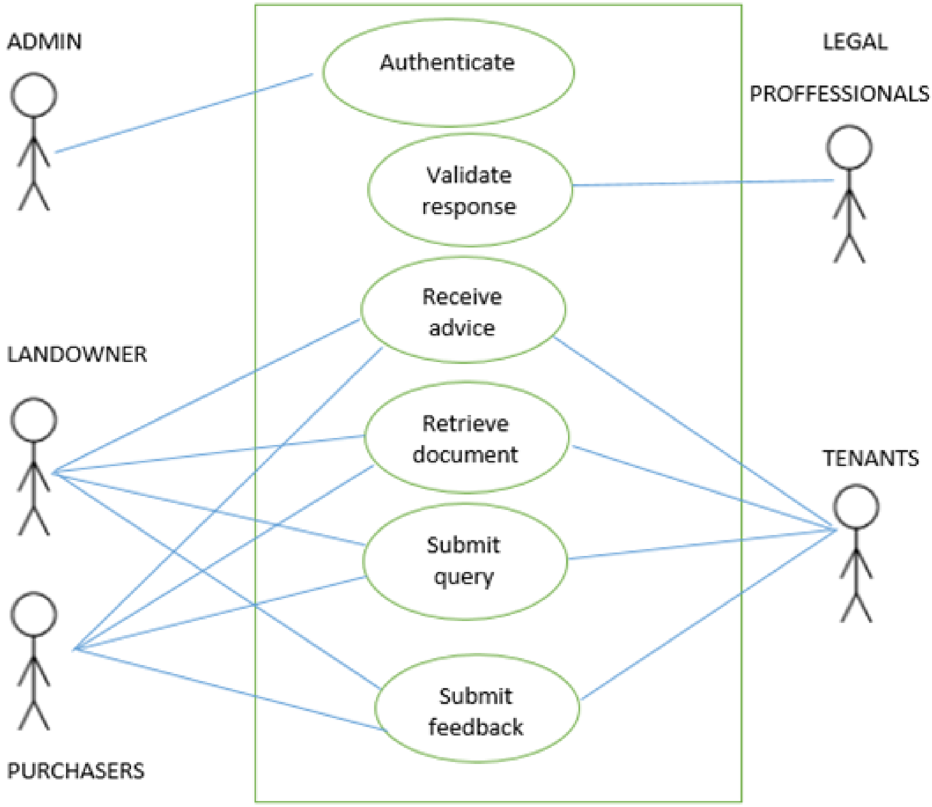

Figure 7.2: Usecase Diagram

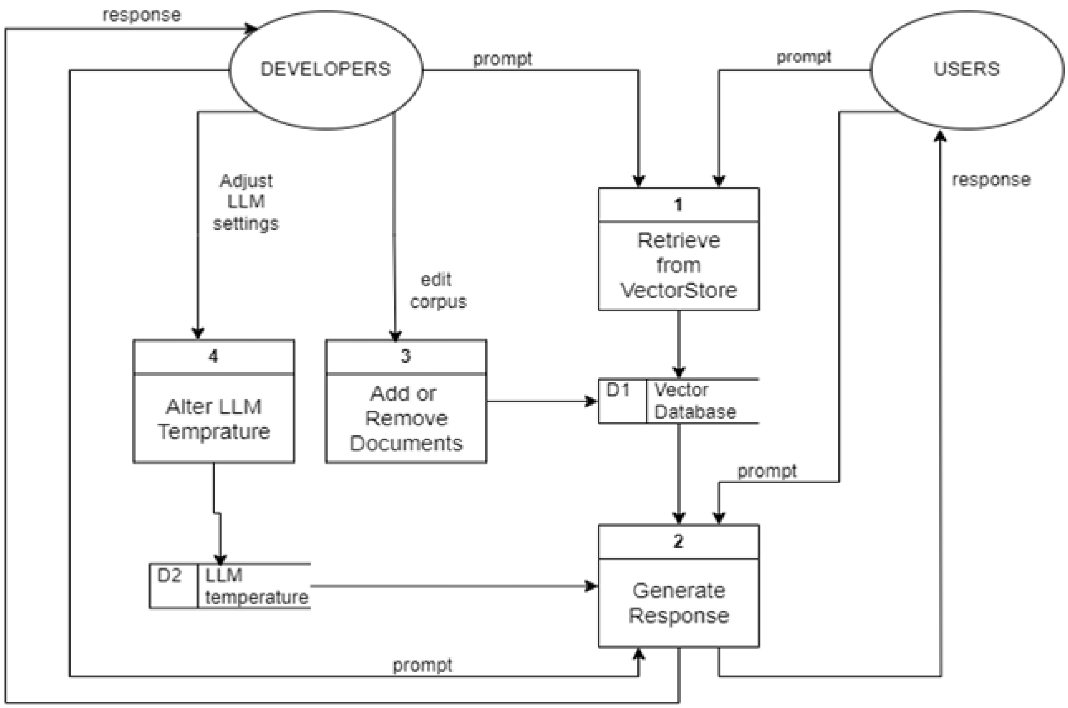

Figure 7.3: Data Flow Diagram