

ĐẠI HỌC QUỐC GIA TP HCM
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN
□ ★ □



Báo cáo Đồ án

Xây Dựng Hệ Thống Hỏi Đáp và So Sánh Luật Đất Đai Dựa Trên Kiến Trúc RAG Tăng Cường bằng Đồ Thị Tri Thức

CS2231.CH191: Mô hình tri thức quan hệ và
ứng dụng

Giảng viên: GVC-TS Nguyễn Đình Hiền

Học viên thực hiện:

Văn Đức Ngo	240101020
Phạm Thăng Long	240101016
Võ Lê Phú Xuân	240101032
Nguyễn Hoàng Hải	240101008

Tóm Tắt

Luật Đất đai là một trong những bộ luật phức tạp và có tầm ảnh hưởng sâu rộng nhất tại Việt Nam. Việc ban hành Luật Đất đai 2024, với nhiều thay đổi mang tính bước ngoặt so với phiên bản 2013, đã tạo ra một thách thức lớn trong việc tiếp cận và áp dụng pháp luật. Ngôn ngữ pháp lý chặt chẽ và khối lượng thông tin lớn đòi hỏi một công cụ tra cứu mạnh mẽ, vượt qua giới hạn của các phương pháp tìm kiếm truyền thống.

Dự án này trình bày việc thiết kế và triển khai một hệ thống Trợ lý Pháp lý Thông minh, nhằm giải quyết thách thức trên bằng cách cung cấp khả năng Hỏi-Đáp và So sánh giữa hai phiên bản luật một cách tự động và đáng tin cậy. Để khắc phục rủi ro ảo giác (hallucination) cố hữu của các Mô hình Ngôn ngữ Lớn (LLM), dự án được xây dựng dựa trên kiến trúc tiên tiến Retrieval-Augmented Generation (RAG).

Điểm cốt lõi của hệ thống nằm ở việc xây dựng một nền tảng tri thức lai (hybrid knowledge base) có cấu trúc kép. Thứ nhất, một Đồ thị Tri thức (Knowledge Graph) bằng Neo4j được xây dựng thông qua phương pháp trích xuất bán tự động có sự hỗ trợ của LLM (Google Gemini), mô hình hóa các thực thể pháp lý cốt lõi (ĐiềuLuật, KháiNiệm, ChủThể...) và các mối quan hệ ngữ nghĩa, đặc biệt là các cạnh liên kết SỬA_ĐỔI_BỒ_SUNG và THAY_THẾ_HOÀN_TOÀN giữa hai phiên bản luật. Thứ hai, một Cơ sở dữ liệu Vector được xây dựng bằng FAISS và mô hình embedding chuyên cho tiếng Việt (bkai-foundation-models/vietnamese-bi-encoder), cho phép tìm kiếm tương đồng ngữ nghĩa hiệu quả.

Pipeline xử lý truy vấn được thiết kế theo mô hình hai giai đoạn (Search-then-Rerank). Giai đoạn đầu, hệ thống truy xuất các điều luật ứng viên. Giai đoạn hai, một mô hình Cross-Encoder (cross-encoder/ms-marco-MiniLM-L-6-v2) được sử dụng để sắp xếp lại (rerank) chính xác các ứng viên này. Cuối cùng, thông qua kỹ thuật Prompt Engineering tiên tiến, kết hợp Chain-of-Thought và các ràng buộc chặt chẽ, LLM được hướng dẫn để sinh ra các câu trả lời chi tiết và có trích dẫn pháp lý rõ ràng. Báo cáo này trình bày chi tiết từng bước trong quy trình xây dựng, từ xử lý dữ liệu đến triển khai, và trình bày kết quả đánh giá định lượng so sánh hiệu năng của hệ thống RAG-KG với một LLM thuần trên bộ 100 câu hỏi thực tế, qua đó chứng minh tính ưu việt của kiến trúc được đề xuất.

1. Giới Thiệu

1.1. Bối cảnh và Động lực Nghiên cứu

Đất đai là nguồn tài nguyên đặc biệt, là tư liệu sản xuất không thể thay thế và có vai trò quyết định đối với sự phát triển kinh tế - xã hội, quốc phòng, an ninh của mỗi quốc gia. Tại Việt Nam, hệ thống pháp luật về đất đai luôn là một trong những lĩnh vực phức tạp và nhận được sự quan tâm lớn nhất từ xã hội. Sự ra đời của Luật Đất đai (sửa đổi) năm 2024, thay thế cho phiên bản năm 2013, đã đánh dấu một bước tiến quan trọng trong việc hoàn thiện thể chế, chính sách, với hàng loạt thay đổi mang tính đột phá về quy hoạch, thu hồi đất, định giá đất, và quyền lợi của người sử dụng đất.

Sự thay đổi mang tính hệ thống này đã tạo ra một khoảng trống thông tin đáng kể. Cả người dân, doanh nghiệp, các nhà đầu tư và thậm chí cả các chuyên gia pháp lý đều đối mặt với nhu cầu cấp thiết trong việc: (1) Nắm bắt nhanh chóng và chính xác các quy định mới; (2) Hiểu rõ bản chất của sự thay đổi so với luật cũ; và (3) Áp dụng các quy định vào các tình huống thực tế. Tuy nhiên, việc tiếp cận văn bản pháp luật gốc gặp nhiều rào cản:

- **Độ phức tạp về ngôn ngữ:** Ngôn ngữ pháp lý sử dụng thuật ngữ chuyên ngành, cấu trúc câu phức tạp, đòi hỏi kiến thức nền tảng để diễn giải đúng.
- **Độ phức tạp về cấu trúc:** Một điều luật có thể tham chiếu đến nhiều điều, khoản, điểm khác trong cùng một văn bản hoặc thậm chí các văn bản luật khác, tạo ra một mạng lưới thông tin chằng chịt.
- **Khối lượng thông tin lớn:** Việc đọc và so sánh thủ công hai bộ luật đồ sộ là một công việc tốn nhiều thời gian và dễ xảy ra sai sót.

Các công cụ tìm kiếm truyền thống dựa trên từ khóa (keyword-based search) tỏ ra không hiệu quả, vì chúng không thể nắm bắt được ngữ nghĩa (ví dụ: chấm dứt quyền sử dụng đất và thu hồi đất có thể liên quan nhưng dùng từ ngữ khác nhau) hay các mối quan hệ logic ẩn sau văn bản.

1.2. Các Thách thức và Cơ hội từ Trí tuệ Nhân tạo

Sự phát triển vượt bậc của các Mô hình Ngôn ngữ Lớn (LLMs) đã mở ra một hướng đi đầy hứa hẹn. Với khả năng hiểu và sinh ngôn ngữ tự nhiên ở mức độ cao, LLMs có tiềm năng trở thành các trợ lý pháp lý ảo, giúp diễn giải và tóm tắt các văn bản phức tạp. Tuy nhiên, việc áp dụng LLM thuần (standalone LLM) trong lĩnh vực pháp lý đối mặt với một thách thức chí mạng: hiện tượng ảo giác. Các mô hình này có thể sáng tạo ra các thông tin không chính xác hoặc không có thật để lấp đầy khoảng trống kiến thức, một rủi ro không thể chấp nhận được khi tính chính xác và căn cứ pháp lý là yêu cầu tối thượng.

Kiến trúc Retrieval-Augmented Generation (RAG) ra đời như một giải pháp nền tảng để giải quyết vấn đề này. Bằng cách tách biệt quá trình truy xuất thông tin (Retrieval) khỏi quá trình sinh câu trả lời (Generation), RAG buộc LLM phải dựa vào một nguồn kiến thức bên ngoài đáng tin cậy. Tuy nhiên, hiệu quả của một hệ thống RAG phụ thuộc trực tiếp vào chất lượng của bước Retrieval. Một hệ thống truy xuất chỉ dựa trên tìm kiếm tương đồng văn bản (dense vector search) vẫn có thể bỏ sót các thông tin liên quan về mặt logic nhưng khác biệt về mặt câu chữ.

1.3. Mục tiêu và Đóng góp của Dự án

Xuất phát từ những thách thức và cơ hội trên, dự án này đặt ra mục tiêu xây dựng một hệ thống Trợ lý Pháp lý Thông minh chuyên sâu về Luật Đất đai 2013 và 2024. Mục tiêu cụ thể bao gồm:

- **Hỏi-Đáp Tình huống:** Cung cấp câu trả lời chính xác, dễ hiểu cho các câu hỏi của người dùng.
- **So sánh Chuyên sâu:** Tự động phân tích, đối chiếu và làm nổi bật các điểm khác biệt cốt lõi giữa hai phiên bản luật.
- **Đảm bảo Tính Tin cậy:** Mọi thông tin đưa ra đều phải có căn cứ, đi kèm với trích dẫn trực tiếp.

Để đạt được các mục tiêu này, đóng góp cốt lõi của dự án là việc đề xuất và triển khai một kiến trúc RAG tăng cường bằng Đồ thị Tri thức (Knowledge Graph-Augmented RAG). Cách tiếp cận này vượt trội hơn các hệ thống RAG thông thường bằng cách không chỉ dựa vào sự tương đồng bề mặt của văn bản mà còn khai thác các mối quan hệ logic có cấu trúc ẩn sâu. Thay vì chỉ dựa vào một cơ sở dữ liệu vector, chúng tôi xây dựng một nền tảng tri thức lai (hybrid knowledge base) bao gồm hai thành phần bổ trợ cho nhau:

- **Đồ thị Tri thức (Neo4j):** Đóng vai trò là "bộ não" lưu trữ tri thức có cấu trúc, cho phép thực hiện các truy vấn logic và suy luận trên các mối quan hệ.
- **Cơ sở dữ liệu Vector (FAISS):** Đóng vai trò là "trí nhớ" cho tri thức phi cấu trúc, phục vụ cho việc tìm kiếm ngữ nghĩa tốc độ cao.

Bằng cách kết hợp hai phương pháp truy xuất từ nền tảng tri thức lai này và tinh chỉnh LLM bằng các kỹ thuật Prompt Engineering tiên tiến, chúng tôi kỳ vọng hệ thống sẽ đạt được độ chính xác và độ tin cậy cao, trở thành một công cụ hữu ích cho cộng đồng.

2. Kiến Trúc và Quy Trình Hệ Thống

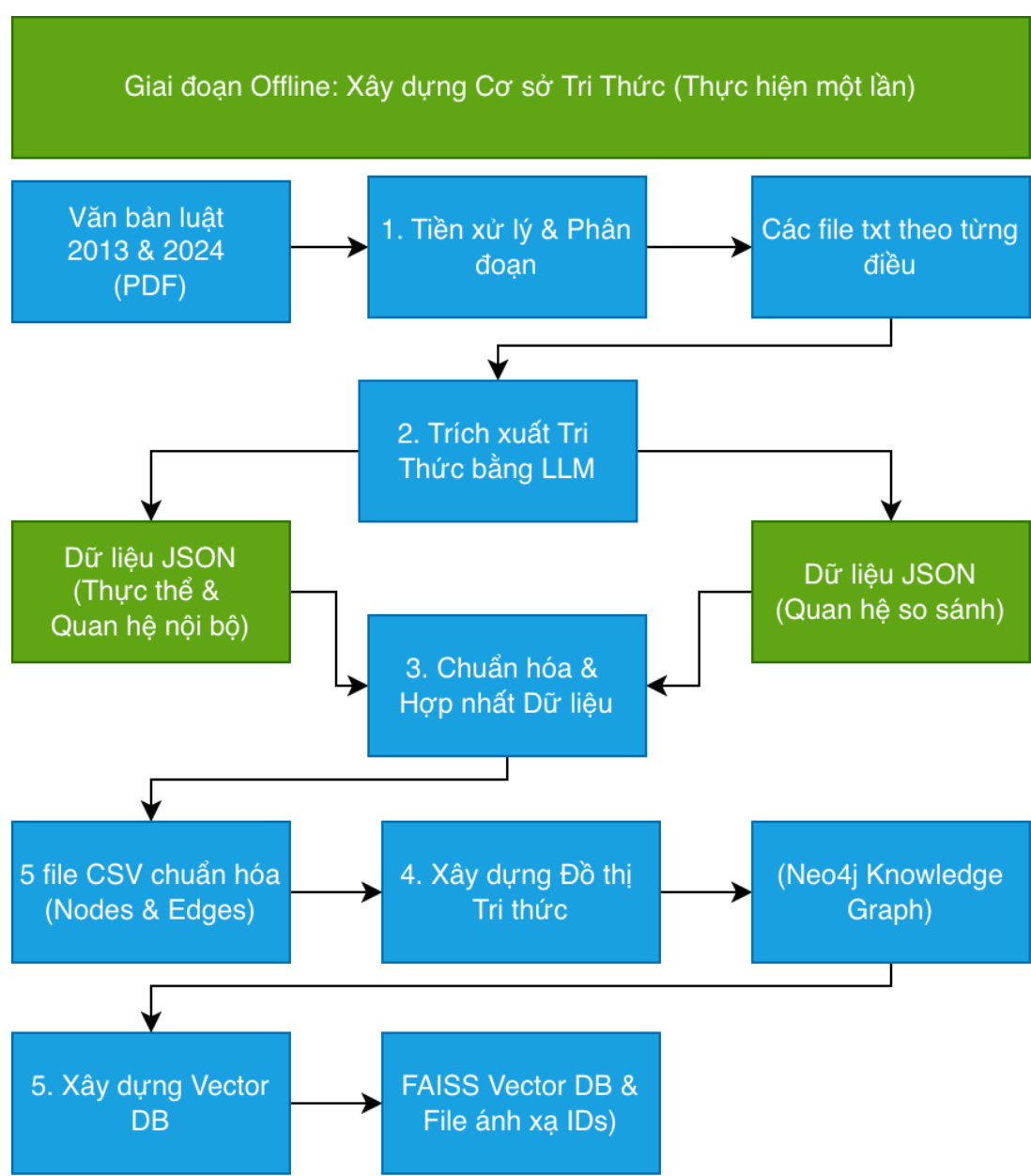
Để xây dựng một Trợ lý Pháp lý có khả năng trả lời chính xác và đáng tin cậy, nhóm đã thiết kế một hệ thống với kiến trúc RAG (Retrieval-Augmented Generation) tiên tiến. Quy trình hoạt động của hệ thống được chia thành hai giai đoạn chính và độc lập:

- **Giai đoạn Offline - Xây dựng Cơ sở Tri thức:** Một quy trình xử lý dữ liệu phức tạp, thực hiện một lần để xây dựng và lập chỉ mục cho toàn bộ tri thức từ các văn bản luật.
- **Giai đoạn Online - Xử lý Truy vấn:** Một pipeline thời gian thực nhận câu hỏi từ người dùng, truy xuất thông tin liên quan từ cơ sở tri thức và sinh ra câu trả lời.

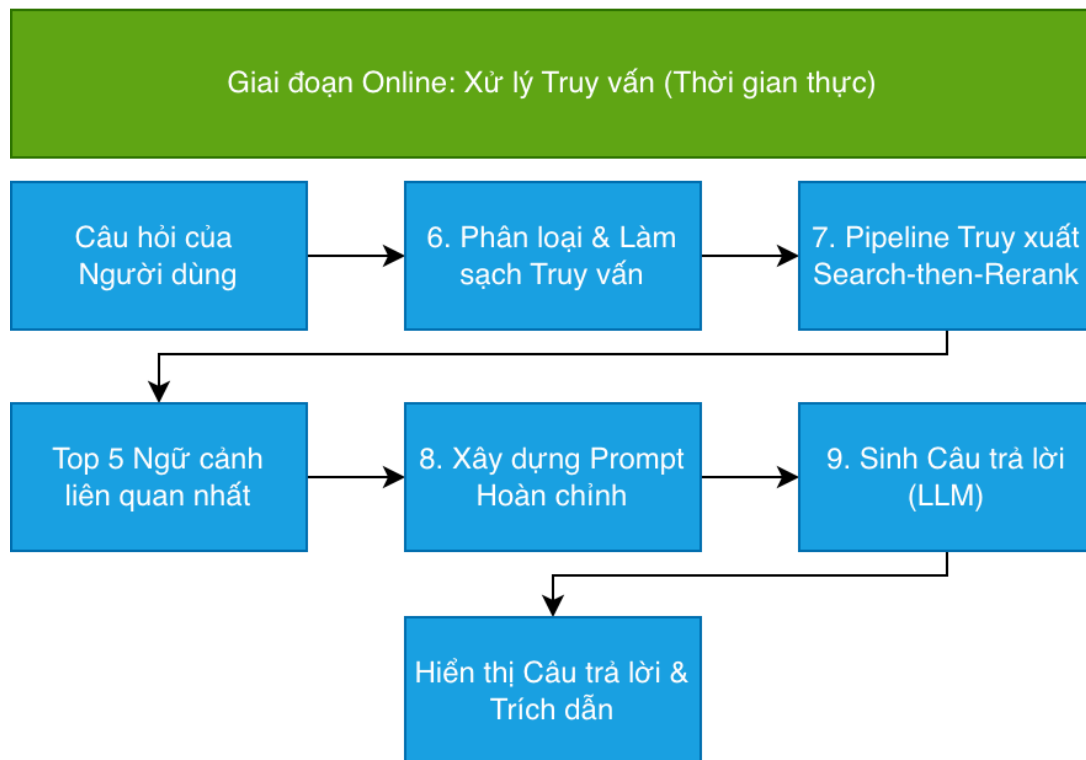
2.1. Sơ đồ luồng Tổng thể

Kiến trúc tổng thể của hệ thống được mô tả trong Sơ đồ 1, bao gồm hai giai đoạn chính hoạt động độc lập:

a) Giai đoạn Offline



b) Giai đoạn Online



2.2. Giai đoạn Offline: Xây dựng Cơ sở Tri thức

Đây là giai đoạn nền tảng, thực hiện một lần để xây dựng toàn bộ "bộ não" cho hệ thống, đảm bảo dữ liệu được xử lý, cấu trúc hóa và lập chỉ mục một cách tối ưu cho việc truy xuất sau này. Quy trình bao gồm các bước tuần tự sau:

- **Bước 1: Tiền xử lý và Phân đoạn (Preprocessing & Chunking)**
 - Input: Hai file PDF của Luật Đất đai 2013 và 2024.
 - Triển khai: Sử dụng thư viện PyPDF2 để trích xuất văn bản, áp dụng Regex để làm sạch, và phân tách theo mẫu Điều \d+\.. để tạo các chunk theo từng Điều luật.
 - Output: Hai thư mục chunks_2013/ và chunks_2024/ chứa các file .txt của từng Điều.
- **Bước 2: Trích xuất Tri thức bằng LLM**
 - Input: Các file .txt từ bước 1.
 - Triển khai: Sử dụng LLM (Google Gemini) với 2 prompt chuyên biệt để trích xuất thực thể, quan hệ nội bộ và quan hệ so sánh. Quá trình được tăng tốc bằng ThreadPoolExecutor.
 - Output: Các thư mục output_json_{year}/ và comparisons_json/ chứa dữ liệu JSON thô.
- **Bước 3: Chuẩn hóa và Hợp nhất Dữ liệu**

- Input: Dữ liệu JSON thô từ bước 2.
- Triển khai: Sử dụng script để tổng hợp các thực thể, hỗ trợ việc tạo một "từ điển đồng nghĩa" (synonym_groups) bán thủ công. Một script chuẩn hóa sau đó áp dụng từ điển này để hợp nhất các nút, chuẩn hóa label, và cập nhật các mối quan hệ.
- Output: 5 file CSV (graph_nodes_2013.csv, graph_nodes_2024.csv, graph_edges_2013.csv, graph_edges_2024.csv, graph_edges_comparison.csv) sẵn sàng cho việc import.
- **Bước 4: Xây dựng Đồ thị Tri thức (KG)**
 - Input: 5 file CSV đã được chuẩn hóa.
 - Triển khai: Sử dụng công cụ neo4j-admin database import full để nạp toàn bộ dữ liệu vào Neo4j. Sau đó, tạo Full-Text Index bằng lệnh Cypher.
 - Output: Một cơ sở dữ liệu đồ thị Neo4j hoàn chỉnh.
- **Bước 5: Xây dựng Cơ sở dữ liệu Vector (Vector DB)**
 - Input: Cơ sở dữ liệu Neo4j.
 - Triển khai: Truy vấn toàn bộ 454 điều luật từ Neo4j. Sử dụng mô hình bkai-foundation-models/vietnamese-bi-encoder để tạo embedding cho từng điều. Lập chỉ mục và lưu trữ các vector bằng thư viện FAISS (IndexFlatIP).
 - Output: Hai file faiss_index.bin và law_ids.json.

Kết thúc Giai đoạn Offline: Từ đầu vào là hai file PDF phi cấu trúc, chúng tôi đã xây dựng thành công một nền tảng tri thức lai (hybrid knowledge base) toàn diện, bao gồm một Đồ thị Tri thức trong Neo4j và một Cơ sở dữ liệu Vector trong FAISS, sẵn sàng cho Giai đoạn Online.

2.3. Giai đoạn Online: Xử lý Truy vấn theo Thời gian thực

Đây là giai đoạn tương tác trực tiếp với người dùng. Khi nhận được một câu hỏi, hệ thống sẽ kích hoạt một pipeline xử lý gồm nhiều bước được thiết kế để tối ưu hóa độ chính xác và độ tin cậy của câu trả lời.

Bước 1: Phân loại Truy vấn và Làm sạch (Query Classification & Cleaning)

- Mục tiêu: Hiểu ý định của người dùng và chuẩn hóa câu hỏi để tối ưu cho các bước tìm kiếm.
- Triển khai:
 - Phân loại (Classification): Một lời gọi LLM đầu tiên được thực hiện với một prompt đơn giản để phân loại câu hỏi của người dùng vào hai loại chính: Hỏi-Đáp hoặc So sánh. Việc xác định sớm ý định này giúp hệ thống lựa chọn prompt sinh câu trả lời phù hợp ở bước cuối cùng.

- Làm sạch (Cleaning): Câu hỏi gốc được đưa qua hàm `clean_query`. Hàm này thực hiện các tác vụ: chuyển đổi sang chữ thường, loại bỏ các từ kích hoạt phổ biến (cho tôi hỏi, là gì...) và các ký tự đặc biệt không mang ngữ nghĩa. Kết quả là một `cleaned_query` súc tích, tập trung vào các thuật ngữ cốt lõi, được tối ưu cho việc tìm kiếm.
- Output:
 - `query_type`: "Hỏi-Đáp" hoặc "So sánh".
 - `cleaned_query`: Chuỗi câu hỏi đã được làm sạch và tối ưu cho tìm kiếm.

Bước 2: Pipeline Truy xuất Hai Giai đoạn (Two-Stage Retrieval Pipeline)

- Mục tiêu: Từ hàng trăm điều luật, tìm ra một vài đoạn văn bản (ngữ cảnh) liên quan và chính xác nhất với câu hỏi của người dùng. Pipeline này là trái tim của kiến trúc RAG, đảm bảo chất lượng của thông tin đầu vào cho LLM.
- Triển khai:
 - Giai đoạn 1 - Thu thập Ứng viên (Candidate Retrieval):
 - Sử dụng lớp `SemanticRetriever`, `cleaned_query` được nhúng thành một vector bằng mô hình Bi-Encoder (`bkai-foundation-models/vietnamese-bi-encoder`).
 - Vector này được dùng để tìm kiếm trong FAISS index, truy xuất ra một danh sách lớn các ứng viên tiềm năng (ví dụ, `initial_k=20`) dựa trên độ tương đồng cosine. Giai đoạn này ưu tiên tốc độ và khả năng gợi nhớ (recall), đảm bảo không bỏ sót các điều luật có khả năng liên quan.
 - Giai đoạn 2 - Sắp xếp lại (Reranking):
 - Nội dung chi tiết của 20 điều luật ứng viên được lấy từ Neo4j.
 - Sử dụng lớp `Reranker`, mô hình Cross-Encoder (`cross-encoder/ms-marco-MiniLM-L-6-v2`) được huy động. Thay vì so sánh vector, mô hình này sẽ xử lý trực tiếp 20 cặp (`cleaned_query`, `document_content`).
 - Cross-Encoder thực hiện một phân tích sâu hơn về sự tương quan ngữ nghĩa và trả về một điểm số chính xác hơn cho từng cặp.
 - Danh sách 20 ứng viên được sắp xếp lại dựa trên điểm số này. `final_k=5` điều luật có điểm số cao nhất được chọn làm ngữ cảnh cuối cùng. Giai đoạn này ưu tiên độ chính xác (precision).
- Output:
 - `final_k` (ví dụ: 5) tài liệu (điều luật) có nội dung và điểm số liên quan cao nhất.

Bước 3: Xây dựng Ngữ cảnh và Sinh Câu trả lời (Context Construction & Generation)

- Mục tiêu: Sử dụng ngữ cảnh đã được chọn lọc để tạo ra một câu trả lời hoàn chỉnh, chính xác và có trích dẫn.
- Triển khai:
 - Xây dựng Ngữ cảnh: Nội dung đầy đủ của 5 điều luật cuối cùng được định dạng thành một chuỗi văn bản duy nhất, có cấu trúc rõ ràng, phân tách từng điều luật và ghi rõ nguồn (Điều X Luật Y).
 - Prompt Engineering: Ngữ cảnh này, cùng với câu hỏi gốc của người dùng (để giữ lại sắc thái ban đầu), được đưa vào một trong hai prompt chuyên biệt (build_qa_prompt hoặc build_comparison_prompt) tùy thuộc vào kết quả phân loại ở Bước 1. Các prompt này được thiết kế với các kỹ thuật như Chain-of-Thought và ràng buộc nghiêm ngặt, hướng dẫn LLM phải suy luận logic, đối chiếu thông tin và bắt buộc trích dẫn nguồn.
 - Sinh Câu trả lời: Prompt hoàn chỉnh được gửi đến LLM (Google Gemini). Mô hình sẽ đọc, hiểu và tổng hợp thông tin chỉ từ ngữ cảnh được cung cấp để tạo ra câu trả lời cuối cùng.
 - Hiển thị: Câu trả lời được hiển thị cho người dùng, kèm theo một khu vực có thể mở rộng (expander) để xem chi tiết nội dung của 5 điều luật đã được sử dụng làm căn cứ.
- Output
 - context_string: Một chuỗi văn bản lớn chứa nội dung của 5 điều luật.
 - final_prompt: Prompt hoàn chỉnh được gửi đến LLM.
 - final_answer: Câu trả lời cuối cùng từ LLM để hiển thị cho người dùng.

Kết luận Giai đoạn Online:

Bằng cách kết hợp một pipeline truy xuất hai giai đoạn mạnh mẽ và kỹ thuật Prompt Engineering tiên tiến, Giai đoạn Online đảm bảo rằng hệ thống không chỉ nhanh chóng mà còn có khả năng đưa ra những câu trả lời sâu sắc, đáng tin cậy, giải quyết hiệu quả các yêu cầu phức tạp của người dùng trong lĩnh vực pháp lý.

3. Các Kỹ Thuật Cốt Lõi: Lựa Chọn, Triển Khai và Thảo Luận

Phần này đi sâu vào việc phân tích lý do lựa chọn (rationale) các công nghệ và thuật toán cốt lõi, những thách thức trong quá trình triển khai và thảo luận về hiệu quả của chúng.

3.1. Mô hình Biểu diễn Tri thức: Tại sao lại là Đồ thị Tri thức Lai?

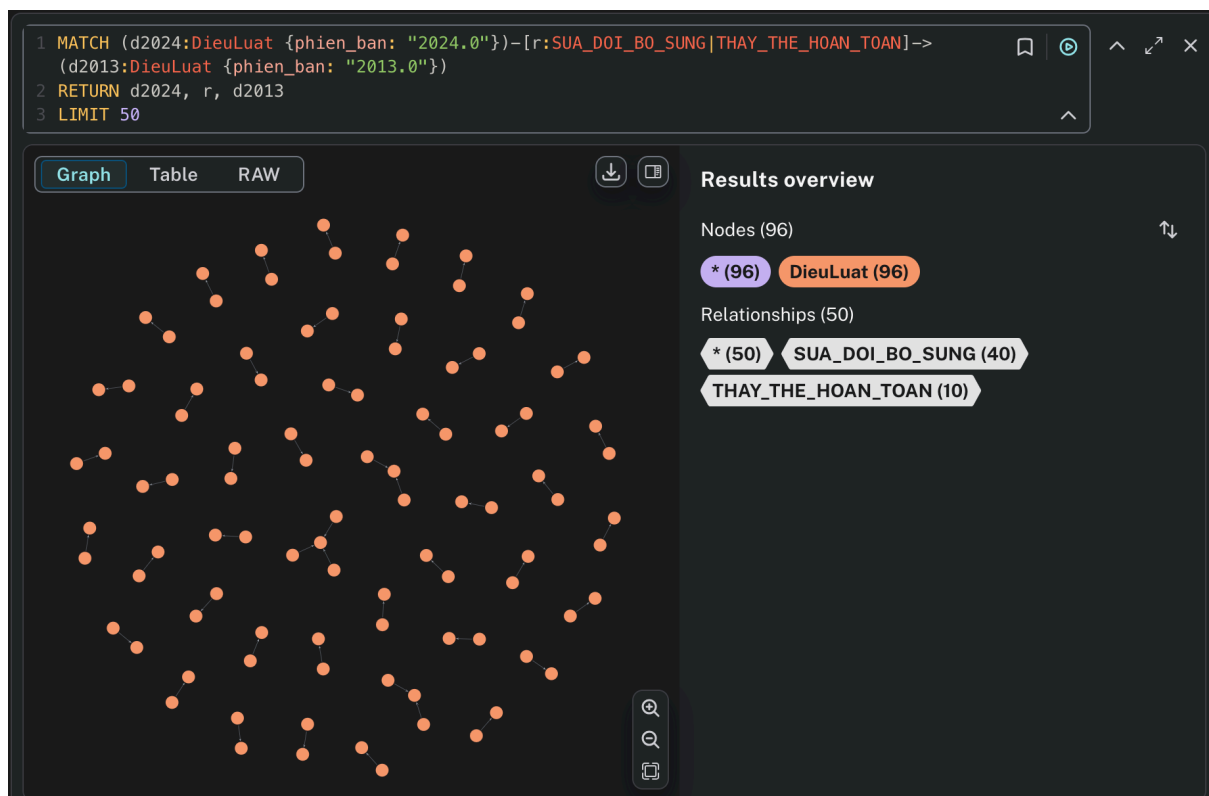
Trong lĩnh vực pháp lý, thông tin tồn tại dưới hai dạng: phi cấu trúc và có cấu trúc. Một hệ thống RAG chỉ dựa trên văn bản thuần túy (Vector DB) sẽ bỏ lỡ các mối quan hệ có cấu trúc

này. Do đó, chúng tôi đã lựa chọn xây dựng một nền tảng tri thức lai (hybrid knowledge base).

- **Đồ thị Tri thức (KG) - Trái tim của sự suy luận:** Chúng tôi chọn Neo4j, một cơ sở dữ liệu đồ thị bản địa (native graph database), làm xương sống cho tri thức có cấu trúc. KG cho phép mô hình hóa các thực thể (Điều luật, Khái niệm,...) như các nút và mối quan hệ (QUY ĐỊNH VỀ, THAY THẾ HOÀN TOÀN) như các cạnh. Điều này cho phép thực hiện các truy vấn suy luận logic phức tạp, một khả năng mà tìm kiếm vector đơn thuần không thể thực hiện được.
- **Cơ sở dữ liệu Vector (Vector DB) - Cánh tay của tìm kiếm ngữ nghĩa:** Chúng tôi chọn FAISS để lưu trữ các vector embedding. FAISS là một thư viện được tối ưu hóa cao cho việc tìm kiếm tương đồng trên các tập vector lớn. Nó cực kỳ hiệu quả về bộ nhớ và tốc độ, phù hợp cho việc tìm kiếm thời gian thực.

Hiệu quả của việc mô hình hóa này được thể hiện rõ qua các hình ảnh trực quan hóa đồ thị.

a) Trực quan hóa Mối quan hệ So sánh giữa Luật Đất đai 2013 và 2024

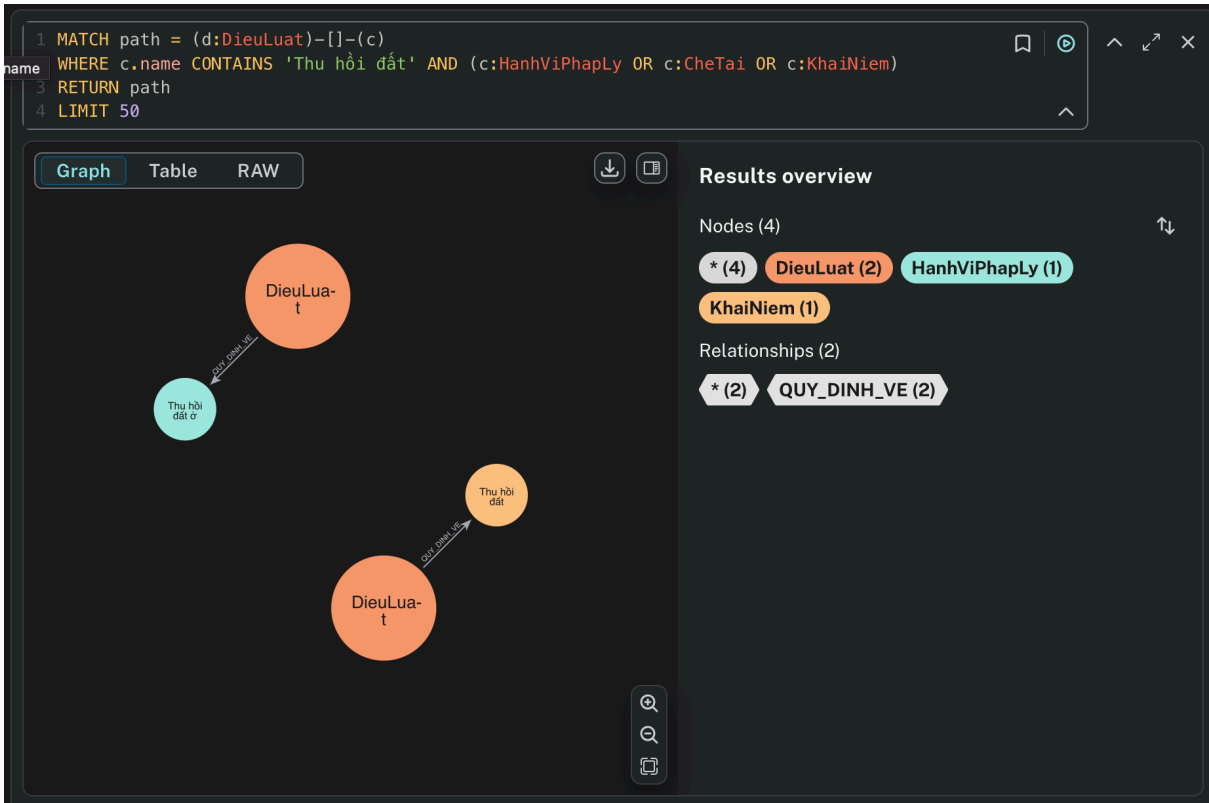


Trực quan hóa Mối quan hệ So sánh giữa Luật Đất đai 2013 và 2024.

Sơ đồ trên là kết quả của truy vấn tìm kiếm 50 cặp điều luật được liên kết bởi các mối quan hệ so sánh. Nó minh họa thành công "cầu nối" tri thức đã được xây dựng giữa hai phiên bản luật. Mỗi cặp nút được kết nối đại diện cho một Điều luật của năm 2024 và một Điều luật tương ứng của năm 2013, thông qua các cạnh SỬA_ĐỔI_BỔ_SUNG hoặc

THAY_THẾ_HOÀN_TOÀN. Cấu trúc này là nền tảng cho phép hệ thống tự động đối chiếu các quy định khi người dùng có yêu cầu so sánh.

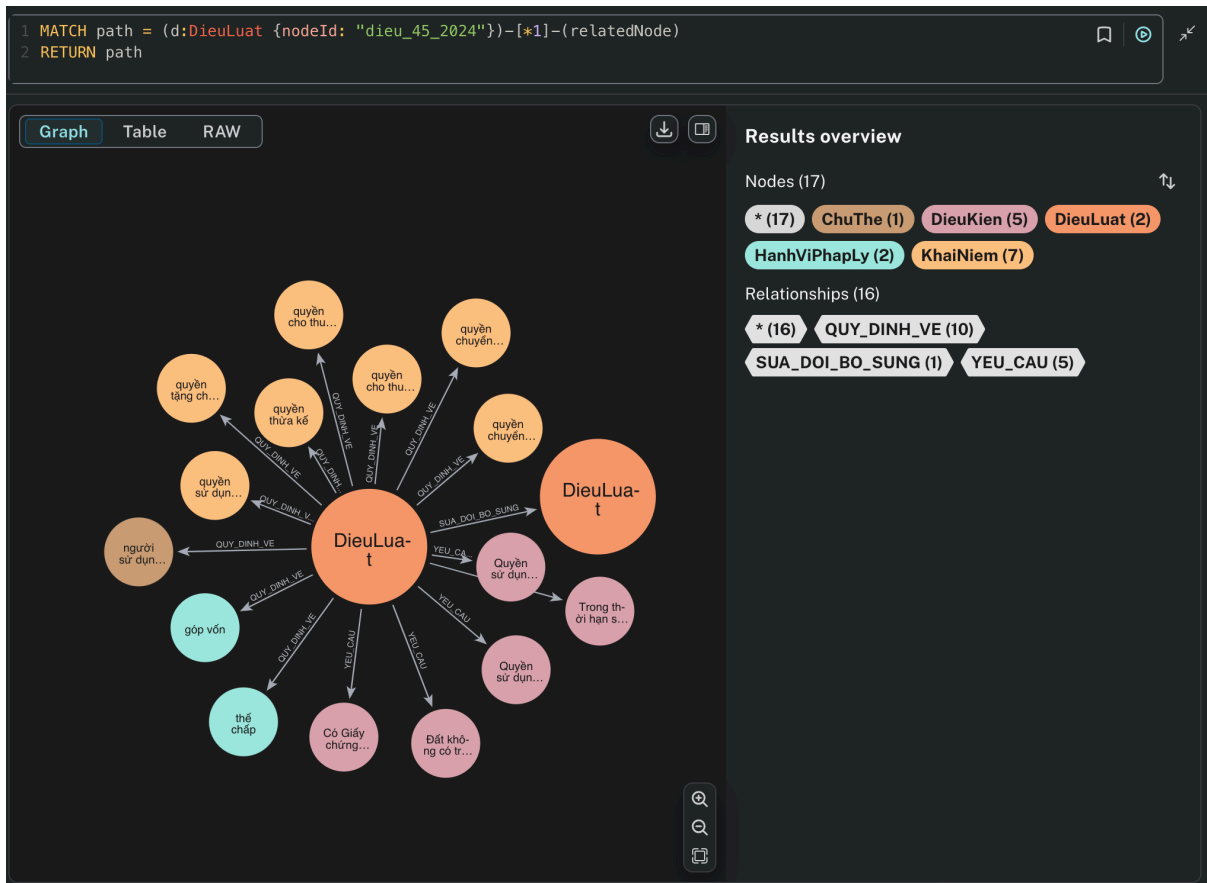
b) Sự hội tụ của các quy định liên quan đến khái niệm "Hộ gia đình"



Sự hội tụ của các quy định liên quan đến khái niệm "Hộ gia đình".

Sơ đồ trên minh họa hiệu quả của quá trình chuẩn hóa và hợp nhất thực thể. Mặc dù khái niệm "Hộ gia đình" được đề cập trong nhiều điều luật khác nhau, hệ thống đã hợp nhất chúng thành một nút trung tâm duy nhất. Từ đó, ta có thể thấy rõ các điều luật liên quan từ cả hai phiên bản luật cùng quy chiếu về một khái niệm cốt lõi. Việc này không chỉ loại bỏ sự dư thừa dữ liệu mà còn cho phép thực hiện các truy vấn phân tích sâu theo chủ đề, ví dụ như "tìm tất cả các quy định liên quan đến 'Hộ gia đình' qua các thời kỳ".

c) Phân rã cấu trúc tri thức của Điều 45 Luật Đất đai 2024



Phân rã cấu trúc tri thức của Điều 45 Luật Đất đai 2024.

Sơ đồ này thể hiện chiều sâu của việc trích xuất tri thức từ một điều luật duy nhất. Nút trung tâm DieuLuat (Điều 45, 2024) không tồn tại một cách cô lập mà được kết nối với một "hệ sinh thái" các thực thể pháp lý mà nó quy định. Như bảng tổng quan cho thấy, điều luật này liên quan trực tiếp đến 17 thực thể khác, bao gồm 7 KháiNiệm (các quyền sử dụng đất), 5 ĐiềuKien cần đáp ứng, và 2 HànhViPhápLy... Cấu trúc chi tiết này cho phép hệ thống hiểu sâu hơn về nội dung của từng điều luật, thay vì chỉ coi nó là một khối văn bản thuần túy, từ đó cung cấp ngữ cảnh chất lượng cao cho LLM.

3.2. Thuật toán Truy xuất: Sự kết hợp giữa Tốc độ và Độ chính xác (Search-then-Rerank)

Một hệ thống RAG tốt phải cân bằng giữa tốc độ và độ chính xác. Để đạt được điều này, chúng tôi đã triển khai pipeline truy xuất hai giai đoạn.

- **Giai đoạn 1: Semantic Search với Bi-Encoder:**
 - Lựa chọn mô hình: Chúng tôi chọn `bkai-foundation-models/vietnamese-bi-encoder`. Đây là một mô hình Bi-Encoder được phát triển bởi BKAI (Đại học Bách Khoa Hà Nội), được huấn luyện chuyên sâu trên dữ liệu tiếng Việt. So với các mô hình đa ngôn

ngữ, nó có khả năng hiểu sắc thái và thuật ngữ tiếng Việt tốt hơn, đặc biệt quan trọng trong lĩnh vực pháp lý. Bi-Encoder có ưu điểm vượt trội về tốc độ vì nó cho phép lập chỉ mục trước toàn bộ tài liệu, nhiệm vụ của nó là tối đa hóa khả năng "gợi nhớ" (recall).

- **Giai đoạn 2: Reranking với Cross-Encoder:**

- Tại sao cần Rerank? Bi-Encoder có thể trả về các kết quả có chung từ khóa nhưng không thực sự khớp với ý định của câu hỏi. Bước Rerank là cần thiết để tinh chỉnh lại thứ hạng.
- Lựa chọn mô hình: Chúng tôi chọn cross-encoder/ms-marco-MiniLM-L-6-v2. Đây là một mô hình Cross-Encoder nhẹ nhưng rất mạnh mẽ, được huấn luyện trên bộ dữ liệu MS MARCO khổng lồ cho các tác vụ xếp hạng. Mặc dù là mô hình đa ngôn ngữ, hiệu năng của nó trong việc so sánh và xếp hạng các cặp câu/văn bản đã được chứng minh rộng rãi. Nó thực hiện một phân tích sâu hơn, tối đa hóa "độ chính xác" (precision).
- Kết quả thực tiễn: Hiệu quả của phương pháp này được thể hiện rõ qua thử nghiệm. Với câu hỏi "khi nào thì nhà nước thu hồi đất do vi phạm pháp luật?", Semantic Search đã trả về Điều 16 (2013) ở vị trí đầu. Tuy nhiên, sau bước Rerank, Điều 81 (2024) (câu trả lời chính xác nhất) đã được đẩy lên vị trí số 1.

3.3. Kỹ thuật Prompt Engineering: Hướng dẫn LLM Suy luận như Chuyên gia

Sau khi đã có ngữ cảnh chất lượng, thách thức tiếp theo là làm thế nào để LLM sử dụng ngữ cảnh đó một cách hiệu quả và đáng tin cậy. Nhóm nhận thấy rằng các prompt đơn giản thường dẫn đến việc LLM trả lời hời hợt hoặc thậm chí suy luận sai (ảo giác) dù đã có ngữ cảnh đúng. Do đó, việc đầu tư vào Prompt Engineering là cực kỳ quan trọng.

- Lựa chọn Kỹ thuật: Các phiên bản prompt cuối cùng của nhóm đã tích hợp nhiều kỹ thuật tiên tiến:
 - Phân vai (Role-playing): Bắt đầu prompt bằng Bạn là một chuyên gia pháp lý đối chiếu văn bản... để đặt LLM vào đúng ngữ cảnh và vai trò cần thực hiện.
 - Suy luận từng bước (Chain-of-Thought - CoT): Thay vì chỉ yêu cầu câu trả lời, nhóm ra lệnh cho LLM tuân theo một QUY TRÌNH PHÂN TÍCH BẮT BUỘC gồm nhiều bước logic (ví dụ: Xác định cặp điều luật -> Đối chiếu song song -> Phân tích điểm khác biệt). Điều này buộc LLM phải suy nghĩ một cách có cấu trúc trước khi trả lời.
 - Ràng buộc Nghiêm ngặt (Strict Constraints): nhóm sử dụng các chỉ dẫn mệnh lệnh và phủ định mạnh mẽ như BẮT BUỘC phải..., TUYỆT ĐỐI KHÔNG

ĐƯỢC kết luận rằng.... Các ràng buộc này có hiệu quả cao trong việc ngăn chặn các hành vi không mong muốn.

- Thảo luận kết quả: Hiệu quả của Prompt Engineering được thể hiện rõ khi so sánh người sử dụng đất. Với prompt ban đầu, LLM đã kết luận sai rằng Luật 2013 không có quy định. Sau khi áp dụng prompt có CoT và ràng buộc nghiêm ngặt, LLM đã đưa ra một bài phân tích đối chiếu chính xác giữa Điều 4 (2024) và Điều 5 (2013). Điều này chứng tỏ rằng, trong các hệ thống RAG, chất lượng của prompt sinh câu trả lời có vai trò quan trọng không kém chất lượng của bước truy xuất.

4. Đánh giá Thực nghiệm và So sánh Hiệu quả

Để đo lường một cách khách quan hiệu quả của kiến trúc RAG tăng cường bằng Đồ thị Tri thức (RAG-KG), chúng tôi đã tiến hành một bài đánh giá định lượng chi tiết. Hiệu năng của hệ thống được so sánh trực tiếp với một Mô hình Ngôn ngữ Lớn (LLM) thuần hoạt động độc lập, nhằm làm nổi bật những ưu điểm và giá trị mà phương pháp của chúng tôi mang lại.

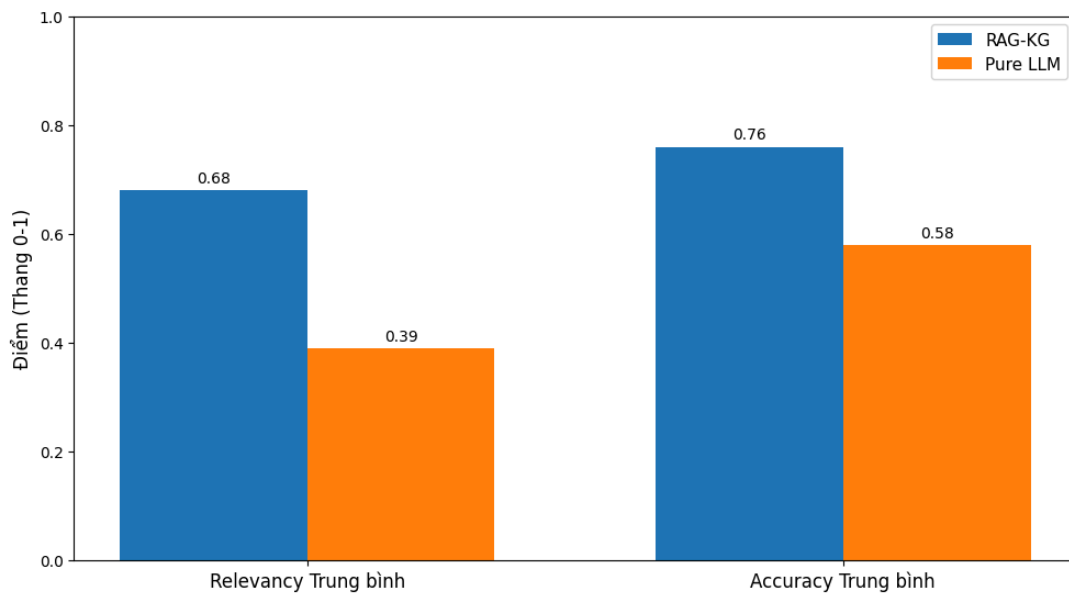
4.1. Thiết lập Thí nghiệm

- **Hệ thống 1 (RAG-KG):** Là hệ thống hoàn chỉnh của dự án, sử dụng Google Gemini làm mô hình sinh ngôn ngữ, được tăng cường bởi nền tảng tri thức lai Neo4j và FAISS.
- **Hệ thống 2 (Pure LLM):** Gửi trực tiếp câu hỏi đến API của Google Gemini mà không có bất kỳ ngữ cảnh nào được cung cấp.
- **Bộ dữ liệu (Test Set):** Một bộ gồm 100 câu hỏi tình huống thực tế do nhóm tự thu thập và biên soạn, bao quát các khía cạnh phức tạp của Luật Đất đai. Bộ dữ liệu được chia thành 5 nhóm chính:
 - Nhóm I: Định nghĩa & Phân loại (10 câu)
 - Nhóm II: Quy hoạch, Kế hoạch sử dụng đất & Thảm quyền (30 câu)
 - Nhóm III: Thu hồi đất, Bồi thường, Hỗ trợ & Tái định cư (30 câu)
 - Nhóm IV: Quyền & Nghĩa vụ của Người sử dụng đất (20 câu)
 - Nhóm V: Hạn mức sử dụng và nhận chuyển quyền (10 câu)
- **Chỉ số đánh giá:**
 - **Answer Relevancy (Độ liên quan):** Đo lường mức độ câu trả lời tập trung vào đúng vấn đề người dùng hỏi, được đánh giá trên thang điểm từ 0 đến 1.
 - **Ground Truth Accuracy (Độ chính xác):** So sánh nội dung của câu trả lời với đáp án chuẩn đã được biên soạn. Thang điểm: 1 (Hoàn toàn đúng), 0.5 (Đúng một phần nhưng chưa đủ hoặc gây nhầm lẫn), 0 (Sai hoàn toàn).

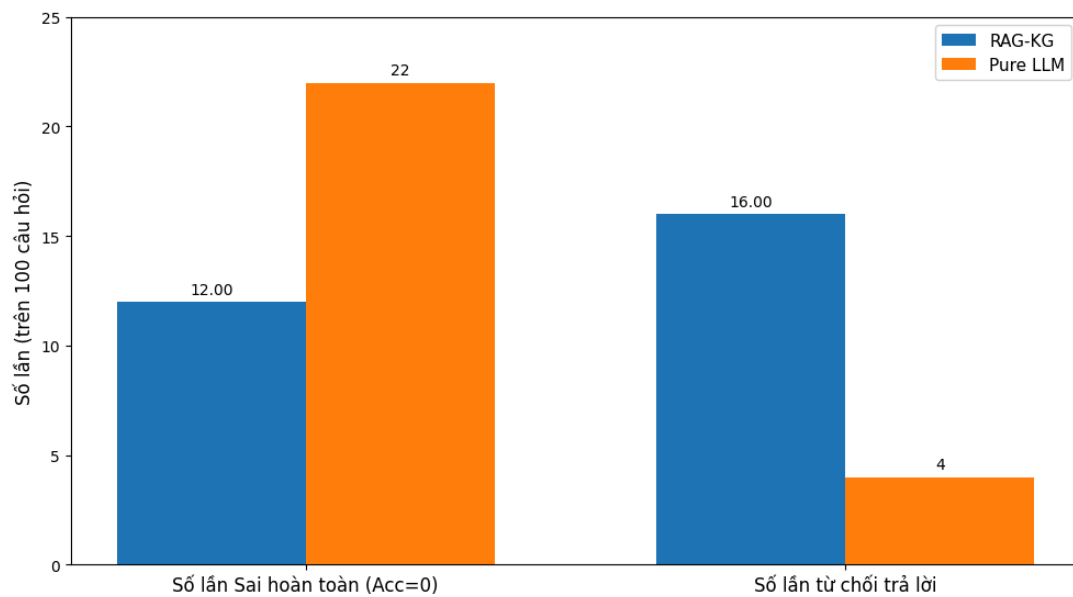
4.2. Kết quả Tổng quan

Hệ thống	Relevancy Trung bình	Accuracy Trung bình	Số lần Sai hoàn toàn (Acc=0)	Số lần từ chối trả lời
RAG-KG	0.678	0.757	12	16
Pure LLM	0.385	0.575	22	4

Biểu đồ 1a: So sánh về Độ liên quan và Độ chính xác



Biểu đồ 1b: So sánh về Hành vi lỗi



Phân tích tổng thể cho thấy hệ thống RAG-KG vượt trội hơn hẳn so với LLM thuần trên mọi phương diện quan trọng:

- **Về Độ liên quan (Relevancy):** Điểm trung bình của RAG-KG (0.68) cao hơn 76% so với LLM thuần (0.39). Điều này chứng tỏ khả năng trả lời tập trung vào đúng vấn đề của hệ thống RAG-KG, trong khi LLM thuần thường đưa ra các câu trả lời lan man, không liên quan.
- **Về Độ chính xác (Accuracy):** RAG-KG đạt điểm chính xác trung bình là 0.76, cao hơn đáng kể so với 0.58 của LLM thuần. Điều này khẳng định việc cung cấp ngữ cảnh từ luật giúp LLM đưa ra câu trả lời đúng đắn và đáng tin cậy hơn.
- **Về Hành vi An toàn và Ảo giác (Hallucination):**
 - LLM thuần trả lời sai hoàn toàn trong 22 trường hợp, cao gần gấp đôi so với 12 trường hợp của RAG-KG.
 - Ngược lại, hệ thống RAG-KG đã chủ động từ chối trả lời trong 16 trường hợp do không tìm được ngữ cảnh phù hợp, trong khi LLM thuần chỉ từ chối 4 lần. Đây không phải là điểm yếu, mà là một hành vi an toàn và có trách nhiệm, thể hiện khả năng "biết mình không biết". LLM thuần có xu hướng "đoán mò" và đưa ra thông tin sai lệch (ví dụ điển hình là các câu II.5, II.14, V.1), một rủi ro nghiêm trọng trong lĩnh vực pháp lý.

4.3. Phân tích Chi tiết theo Từng loại câu hỏi

a) Đối với các câu hỏi Định nghĩa & So sánh (Nhóm I, V):

- LLM thuần thể hiện khả năng "ghi nhớ" tốt các định nghĩa đơn giản, nhưng hoàn toàn thất bại trong việc so sánh (ví dụ, câu V.1 về hạn mức, LLM thuần đã ảo giác rằng "Luật 2024 bỏ quy định"). Ngược lại, hệ thống RAG-KG, nhờ được cung cấp ngữ cảnh từ cả hai bộ luật, đã thực hiện so sánh chính xác và chi tiết, chỉ ra rõ sự thay đổi từ "10 lần" lên "15 lần".

b) Đối với các câu hỏi Tình huống & Thủ tục (Nhóm II, III, IV):

- Đây là lĩnh vực mà hệ thống RAG-KG thể hiện sự vượt trội rõ rệt nhất. Với khả năng truy xuất các điều luật cụ thể, hệ thống đã trả lời đúng phần lớn các câu hỏi phức tạp về thẩm quyền, điều kiện và thời hạn.
- Một ví dụ điển hình là câu II.5 ("Kỳ quy hoạch sử dụng đất cấp huyện theo LĐĐ 2024 là bao nhiêu năm?"). Hệ thống RAG-KG đã truy xuất đúng Điều 62 (2024) và trả lời chính xác là "10 năm". Trong khi đó, LLM thuần đã "đoán mò" và trả lời sai là "5 năm". Sai sót về các chi tiết cụ thể như con số, thẩm quyền, thời hạn là lỗi phổ biến của LLM thuần và cực kỳ nguy hiểm trong bối cảnh tư vấn pháp lý.

4.4. Thảo luận

Kết quả đánh giá định lượng đã khẳng định mạnh mẽ luận điểm của dự án: một kiến trúc RAG được thiết kế tốt, tăng cường bằng Đồ thị Tri thức và pipeline truy xuất hai giai đoạn, có hiệu năng vượt trội so với việc sử dụng LLM một cách độc lập. Việc cung cấp một ngữ cảnh pháp lý chính xác, được chọn lọc kỹ càng, không chỉ giúp tăng độ chính xác mà còn là một cơ chế hiệu quả để kiểm chế hiện tượng "ảo giác", buộc LLM phải đưa ra câu trả lời dựa trên căn cứ thực tế.

Mặc dù hệ thống RAG-KG vẫn có những lần thất bại (chủ yếu do bước truy xuất chưa tìm được điều luật phù hợp), hành vi từ chối trả lời trong những trường hợp này được đánh giá là an toàn và phù hợp hơn so với việc đưa ra thông tin sai lệch.

5. Kết Luận và Hướng Phát Triển Tương Lai

5.1. Kết Luận

Dự án đã thiết kế và triển khai thành công một nguyên mẫu hoạt động của hệ thống Trợ lý Pháp lý Thông minh, giải quyết hiệu quả bài toán hỏi-đáp và so sánh hai phiên bản Luật Đất đai. Bằng cách áp dụng một kiến trúc RAG nâng cao, dự án đã chứng minh được tính khả thi và hiệu quả của việc kết hợp các công nghệ AI tiên tiến vào lĩnh vực pháp lý phức tạp.

Những đóng góp chính của dự án bao gồm:

- Xây dựng thành công một pipeline bán tự động để chuyển đổi văn bản pháp lý phi cấu trúc thành Đồ thị Tri thức (KG). Quy trình này, kết hợp giữa sức mạnh của LLM trong việc trích xuất và các script chuẩn hóa dữ liệu, đã cho thấy tiềm năng to lớn trong việc giảm thiểu công sức thủ công và xây dựng các cơ sở tri thức có cấu trúc một cách nhanh chóng.
- Triển khai và xác thực hiệu quả của pipeline truy xuất hai giai đoạn (Search-then-Rerank). Hệ thống đã cân bằng thành công giữa tốc độ của Bi-Encoder (Semantic Search) và độ chính xác của Cross-Encoder (Reranker), đảm bảo các ngữ cảnh được cung cấp cho LLM có độ liên quan cao nhất, qua đó trực tiếp nâng cao chất lượng câu trả lời.
- Khẳng định vai trò quyết định của Kỹ thuật Prompt Engineering. Các thử nghiệm đã chứng minh rằng, ngay cả khi có cùng một ngữ cảnh chất lượng, việc thiết kế prompt với các chỉ dẫn suy luận từng bước (Chain-of-Thought) và các ràng buộc nghiêm ngặt là yếu tố then chốt để điều khiển hành vi của LLM, ngăn chặn ảo giác và đảm bảo câu trả lời được sinh ra là chính xác, có cấu trúc và đáng tin cậy.

Hệ thống hiện tại có khả năng cung cấp các câu trả lời có căn cứ và các bài phân tích so sánh sâu sắc, trở thành một công cụ hữu ích cho việc tra cứu và tìm hiểu pháp luật đất đai.

5.2. Hạn Chế của Hệ Thống

Mặc dù đã đạt được những kết quả ban đầu khả quan, hệ thống hiện tại vẫn còn một số hạn chế cần được ghi nhận:

- Phụ thuộc vào chất lượng Chunking: Hiệu quả tìm kiếm các định nghĩa cụ thể bị ảnh hưởng bởi việc phân đoạn theo Điều luật. Các điều luật dài chứa nhiều khái niệm (như Điều 3 - Giải thích từ ngữ) có thể làm giảm độ chính xác của semantic search đối với các câu hỏi ngắn.
- Giới hạn Tri thức: Cơ sở tri thức hiện tại chỉ bao gồm hai văn bản luật gốc. Hệ thống chưa có khả năng trả lời các câu hỏi đòi hỏi kiến thức từ các văn bản dưới luật như Nghị định, Thông tư hướng dẫn thi hành.
- Hạn chế của Truy xuất Lai: Pipeline hiện tại, dù hiệu quả, vẫn chủ yếu dựa vào semantic search và reranking. Tiềm năng của Đồ thị Tri thức trong việc thực hiện các suy luận logic phức tạp (ví dụ: "Nếu một cá nhân thuộc đối tượng A, họ sẽ có những quyền gì theo Điều B và Điều C?") chưa được khai thác triệt để trong pipeline một cách tự động.
- Sự phụ thuộc vào chất lượng Prompt: Mặc dù đã được tinh chỉnh, hiệu quả của hệ thống vẫn phụ thuộc nhiều vào khả năng của LLM trong việc tuân thủ các chỉ dẫn phức tạp trong prompt. Việc thay đổi mô hình LLM hoặc phiên bản của nó có thể đòi hỏi phải hiệu chỉnh lại prompt.

5.3. Hướng Phát Triển Tương Lai

Dựa trên những kết quả đã đạt được và các hạn chế còn tồn tại, nhóm đề ra các hướng phát triển chính trong tương lai:

- Đánh giá Định lượng và So sánh Hiệu quả:
 - Mục tiêu: Thực hiện một bài đánh giá khoa học để đo lường hiệu quả của hệ thống.
 - Phương pháp: Sử dụng bộ 100 câu hỏi tình huống đã thu thập làm bộ dữ liệu kiểm thử (test set). Đánh giá và so sánh hiệu năng giữa: (1) Hệ thống RAG-KG hoàn chỉnh của chúng tôi; (2) LLM thuần (Gemini, GPT-4 không có RAG); và (3) Một hệ thống RAG cơ bản (chỉ dùng semantic search). Các chỉ số đánh giá sẽ bao gồm: Faithfulness (độ trung thực của câu trả lời so với nguồn), Answer Relevancy (độ liên quan của câu trả lời với câu hỏi), và Context Precision/Recall (độ chính xác của bước truy xuất).
- Mở rộng và Làm giàu Cơ sở Tri thức:

- Tích hợp Văn bản Dưới luật: Phân tích và tích hợp các Nghị định, Thông tư quan trọng liên quan đến Luật Đất đai. Điều này sẽ yêu cầu mở rộng schema của KG để biểu diễn mối quan hệ giữa Luật và các văn bản hướng dẫn.
- Chunking Thích ứng (Adaptive Chunking): Xây dựng một chiến lược phân đoạn thông minh hơn. Thay vì chỉ chunk theo Điều, hệ thống có thể tự động chunk nhỏ hơn (theo Khoản, Điểm) đối với các điều luật dài và chứa nhiều định nghĩa.
- Tối ưu Pipeline Truy xuất Lai (Hybrid Retrieval):
 - Tích hợp Truy vấn Đồ thị: Xây dựng một module có khả năng phân tích câu hỏi của người dùng và tự động sinh ra các truy vấn Cypher để khai thác các mối quan hệ trong KG.
 - Kết hợp Điểm số: Phát triển một thuật toán để kết hợp một cách thông minh điểm số từ Semantic Search, Full-Text Search và kết quả từ truy vấn đồ thị, nhằm tạo ra một bảng xếp hạng kết quả cuối cùng toàn diện nhất trước khi đưa vào bước Reranking.
- Cải thiện Tương tác Người dùng:
 - Trực quan hóa Đồ thị: Tích hợp một thành phần giao diện cho phép người dùng nhấp vào một điều luật hoặc một khái niệm và xem trực quan các mối quan hệ của nó trong đồ thị.
 - Phản hồi của Người dùng: Xây dựng cơ chế cho phép người dùng đánh giá câu trả lời (hữu ích/không hữu ích), từ đó thu thập dữ liệu để tiếp tục tinh chỉnh hệ thống.

Phụ lục A: Hướng dẫn Cài đặt và Chạy

Toàn bộ mã nguồn, dữ liệu đã xử lý và tài liệu hướng dẫn chi tiết của dự án được công khai tại kho lưu trữ GitHub. Người dùng có thể lựa chọn giữa việc chạy ứng dụng với dữ liệu có sẵn hoặc tự xây dựng lại toàn bộ cơ sở tri thức từ đầu.

A. Yêu Cầu Hệ Thống

- Phần mềm: Python 3.10+, Conda, Java JDK 17+, Neo4j Desktop.
- Phần cứng: Tối thiểu 8GB RAM. GPU (NVIDIA, hỗ trợ CUDA) được khuyến nghị để tăng tốc quá trình xây dựng Vector DB và Reranking.
- Khóa API: Cần có API Key của Google Gemini, được cấu hình trong file `.env`.

B. Kho Lưu Trữ Mã Nguồn (Source Code Repository)

Toàn bộ dự án, bao gồm mã nguồn, dữ liệu mẫu và file README.md với hướng dẫn chi tiết từng bước, có thể được truy cập tại địa chỉ sau:

https://github.com/vanducngo/CS2231.CH191_RelationalKnowledgeModel

C. Hướng dẫn Nhanh

Chi tiết về việc cài đặt môi trường và các lệnh thực thi được mô tả đầy đủ trong file README.md. Dưới đây là tóm tắt quy trình chính:

- Thiết lập môi trường: Cài đặt các thư viện cần thiết bằng Conda và pip theo hướng dẫn trong README.md.
- Nạp Đồ thị Tri thức: Sử dụng công cụ neo4j-admin database import để nạp các file nodes_final.csv và relationships_final.csv vào cơ sở dữ liệu Neo4j.
- Tạo Index: Chạy một lệnh Cypher đơn giản trong Neo4j Browser để tạo Full-Text Index.
- Khởi động: Chạy lệnh streamlit run app.py để khởi động giao diện web.

D. Để Xây dựng lại Cơ sở Tri thức từ đầu:

Quy trình này bao gồm việc chạy một chuỗi các script Python theo thứ tự được chỉ định trong README.md. Chuỗi script này sẽ tự động hóa toàn bộ các bước, từ việc xử lý file PDF gốc, trích xuất tri thức bằng LLM, chuẩn hóa dữ liệu, đến việc tạo ra các file CSV cuối cùng và xây dựng Vector DB (FAISS index).

Nhóm khuyến khích tham khảo file README.md tại Source Code trên github để có được hướng dẫn chi tiết và cập nhật nhất.

Phụ lục B: Giao diện và Ví dụ Hoạt động của Ứng dụng

Phần này trình bày các hình ảnh minh họa giao diện người dùng (UI) của hệ thống Trợ lý Pháp lý, được xây dựng bằng Streamlit. Các ví dụ thể hiện hai chức năng cốt lõi: So sánh Luật và Hỏi-Đáp Tình huống.

B.1. Chức năng So sánh Luật

Chức năng này cho phép người dùng nhập một chủ đề và hệ thống sẽ tự động đối chiếu các quy định liên quan giữa hai phiên bản Luật Đất đai 2013 và 2024.

a) Giao diện chức năng So sánh Luật.

Người dùng nhập yêu cầu so sánh định nghĩa "người sử dụng đất" và gửi đi.

Hỏi đáp, tra cứu và so sánh về Luật Đất đai 2013 và 2024.

Hỏi-Đáp & Tra cứu

So sánh Luật

So sánh sự khác biệt giữa Luật 2013 và 2024

Nhập chủ đề bạn muốn so sánh:

So sánh định nghĩa về "người sử dụng đất" giữa Luật Đất đai 2013 và 2024

So sánh

Đang tìm kiếm cho: 'so sánh về người sử dụng đất giữa luật đất đai 2013 và 2024'

b) Kết quả phân tích so sánh.

Hệ thống trả về một bài phân tích chi tiết, có cấu trúc, làm nổi bật các điểm khác biệt cốt lõi giữa hai bộ luật dựa trên ngữ cảnh đã được truy xuất.

Bài phân tích so sánh:

So sánh định nghĩa về "người sử dụng đất" giữa Luật Đất đai 2013 và 2024

Việc so sánh định nghĩa "người sử dụng đất" giữa Luật Đất đai 2013 và 2024 cần lưu ý rằng, Luật Đất đai 2013 không có một điều khoản cụ thể định nghĩa trực tiếp khái niệm "người sử dụng đất". Thay vào đó, nó được xác định thông qua các quy định về đối tượng được Nhà nước giao đất, cho thuê đất, công nhận quyền sử dụng đất. Ngược lại, Luật Đất đai 2024 (Điều 4) đã đưa ra một định nghĩa rõ ràng hơn về "người sử dụng đất".

Dưới đây là phân tích chi tiết:

Luật Đất đai 2013:

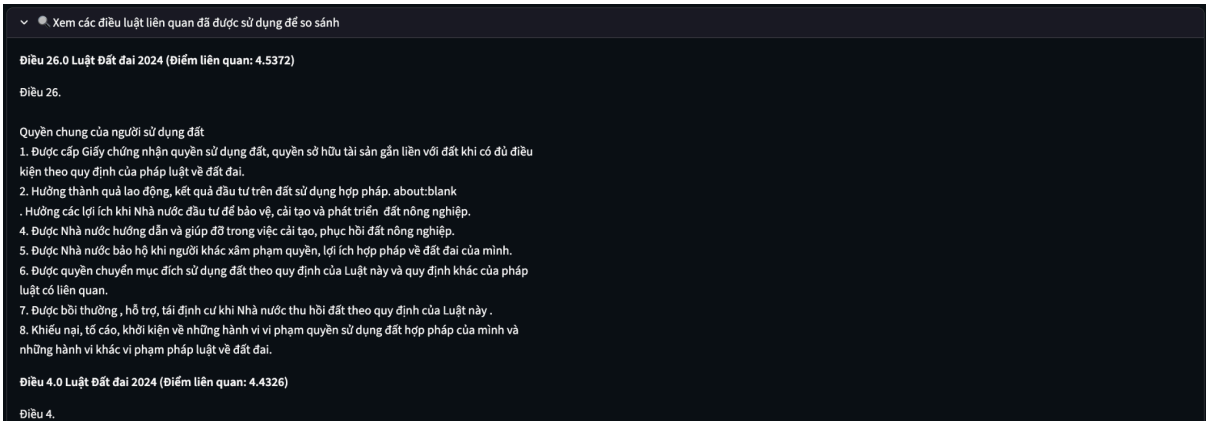
- Không có định nghĩa rõ ràng: Luật Đất đai 2013 không có một điều khoản cụ thể định nghĩa "người sử dụng đất".
- Xác định gián tiếp: Khái niệm "người sử dụng đất" được hiểu thông qua các quy định về các chủ thể được Nhà nước giao đất, cho thuê đất, công nhận quyền sử dụng đất, nhận chuyển quyền sử dụng đất.
- Điều 4 (Sở hữu đất đai): Đề cập đến việc "Nhà nước trao quyền sử dụng đất cho người sử dụng đất theo quy định của Luật này." Tuy nhiên, điều này chỉ khẳng định sự tồn tại của "người sử dụng đất" chứ không định nghĩa họ là ai.
- Điều 5 (Người sử dụng đất): Liệt kê các đối tượng được giao đất, cho thuê đất, công nhận quyền sử dụng đất, nhận chuyển quyền sử dụng đất như tổ chức, hộ gia đình, cá nhân, cộng đồng dân cư, cơ sở tôn giáo, tổ chức nước ngoài, người Việt Nam định cư ở nước ngoài, tổ chức kinh tế có vốn đầu tư nước ngoài. Đây là cách mà Luật Đất đai 2013 gián tiếp xác định ai là người sử dụng đất.
- Trọng tâm: Tập trung vào việc liệt kê các chủ thể có quyền sử dụng đất.

Luật Đất đai 2024 (Điều 4):

- Định nghĩa trực tiếp: "Người sử dụng đất được Nhà nước giao đất, cho thuê đất, công nhận quyền sử dụng đất; đang sử dụng đất ổn định, đủ điều kiện cấp Giấy chứng nhận quyền sử dụng đất, quyền sở hữu tài sản gắn liền với đất mà chưa được Nhà nước cấp Giấy chứng nhận quyền sử dụng đất, Giấy chứng nhận quyền sở hữu nhà ở và quyền sử dụng đất ở, Giấy chứng nhận quyền sử dụng đất, quyền sở hữu nhà ở và tài sản khác gắn liền với đất, Giấy chứng nhận quyền sử dụng đất, quyền sở hữu tài sản gắn liền với đất; nhận quyền sử dụng đất; thuê lại đất theo quy định của Luật này".
- Liệt kê chi tiết các đối tượng: Tiếp tục liệt kê chi tiết các đối tượng là người sử dụng đất, bao gồm:
 - Tổ chức trong nước (cơ quan nhà nước, tổ chức chính trị - xã hội, tổ chức kinh tế, v.v.)
 - Tổ chức tôn giáo
 - Cá nhân trong nước, người Việt Nam định cư ở nước ngoài

c) Ngữ cảnh truy xuất cho chức năng So sánh.

Đây là một tính năng quan trọng, thể hiện sự minh bạch của hệ thống. Người dùng có thể mở rộng để xem chi tiết nội dung của 5 điều luật liên quan nhất (sau khi đã qua bước Reranking) đã được sử dụng làm cơ sở để LLM tạo ra bài phân tích.

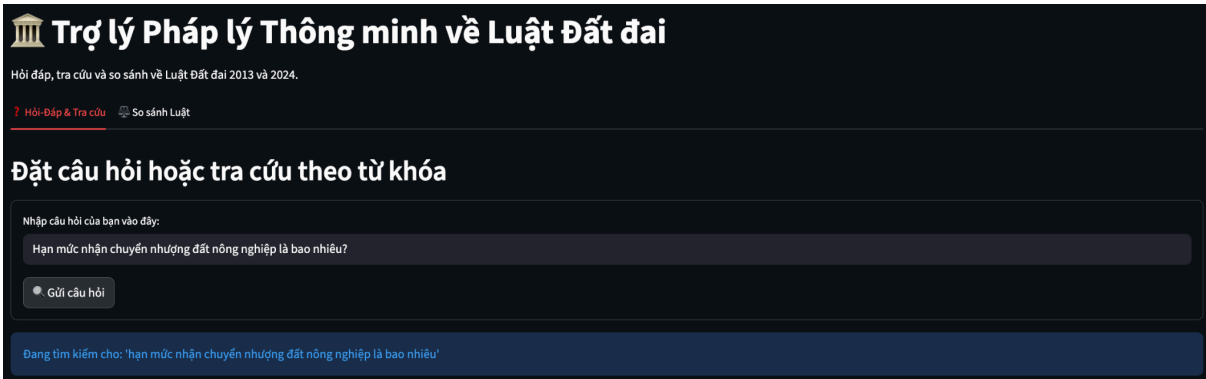


B.2. Chức năng Hỏi-Đáp Tình huống

Chức năng này cho phép người dùng đặt các câu hỏi bằng ngôn ngữ tự nhiên và nhận lại câu trả lời có căn cứ pháp lý.

a) Giao diện chức năng Hỏi-Đáp.


Người dùng nhập một câu hỏi tình huống cụ thể về "hạn mức nhận chuyển nhượng đất nông nghiệp".



b) Kết quả trả lời và cơ chế Phản hồi.

Hệ thống cung cấp một câu trả lời chi tiết, tổng hợp thông tin từ nhiều điều luật liên quan. Điểm nhấn mạnh về tính minh bạch và tương tác được thể hiện qua:

- Cơ chế Phản hồi:** Các nút "👍 Hữu ích" và "👎 Không hữu ích" cho phép người dùng đánh giá chất lượng câu trả lời, tạo ra một kênh thu thập dữ liệu quý giá để cải thiện hệ thống trong tương lai.
- Trích dẫn trong câu trả lời:** Câu trả lời của LLM được huấn luyện để luôn đi kèm với căn cứ pháp lý rõ ràng.

**Câu trả lời:**

Để trả lời câu hỏi về hạn mức nhận chuyển nhượng đất nông nghiệp, chúng ta cần xem xét cả Luật Đất đai 2024 và Luật Đất đai 2013, vì các quy định này có sự khác biệt và có thể áp dụng tùy theo thời điểm giao dịch.

1. Đối với cá nhân:

- Luật Đất đai 2024 (Điều 177):** Hạn mức nhận chuyển quyền sử dụng đất nông nghiệp của cá nhân *không quá 15 lần* hạn mức giao đất nông nghiệp của cá nhân đối với mỗi loại đất quy định tại các khoản 1, 2 và 3 Điều 176 của Luật này.
- Luật Đất đai 2013 (Điều 130):** Hạn mức nhận chuyển quyền sử dụng đất nông nghiệp của hộ gia đình, cá nhân *không quá 10 lần* hạn mức giao đất nông nghiệp của hộ gia đình, cá nhân đối với mỗi loại đất quy định tại các khoản 1, 2 và 3 Điều 129 của Luật này.

Như vậy, Luật Đất đai 2024 đã nâng hạn mức nhận chuyển quyền sử dụng đất nông nghiệp của cá nhân lên 15 lần so với 10 lần theo Luật Đất đai 2013.

Lưu ý quan trọng:


- Hạn mức cụ thể sẽ phụ thuộc vào quy định của Ủy ban nhân dân cấp tỉnh, căn cứ vào điều kiện cụ thể của địa phương (theo Điều 177.3 của Luật Đất đai 2024). Do đó, cần tham khảo quy định cụ thể của từng tỉnh để biết chính xác hạn mức được phép.
- Hạn mức này áp dụng cho *từng loại* đất nông nghiệp (ví dụ: đất trồng lúa, đất trồng cây hàng năm khác, đất trồng cây lâu năm...).


2. Đối với tổ chức kinh tế, hộ gia đình (khi thực hiện dự án đầu tư sản xuất, kinh doanh phi nông nghiệp):

- Luật Đất đai 2013 (Điều 193):** Không quy định hạn mức cụ thể mà quy định các điều kiện để được nhận chuyển nhượng, nhận góp vốn, thuê quyền sử dụng đất nông nghiệp, bao gồm:
 - Có văn bản chấp thuận của cơ quan nhà nước có thẩm quyền.
 - Mục đích sử dụng đất phù hợp với quy hoạch, kế hoạch sử dụng đất.
 - Tuân thủ quy định về đất chuyển trồng lúa nước (nếu có).

Kết luận:


Hạn mức nhận chuyển nhượng đất nông nghiệp cho cá nhân phụ thuộc vào Luật Đất đai có hiệu lực tại thời điểm giao dịch, cụ thể là 10 lần theo Luật Đất đai 2013 và 15 lần theo Luật Đất đai 2024, đồng thời phải tuân theo quy định cụ thể của UBND cấp tỉnh. Đối với tổ chức kinh tế, hộ gia đình nhận chuyển nhượng để thực hiện dự án đầu tư sản xuất, kinh doanh phi nông nghiệp, không có hạn mức cụ thể mà phải đáp ứng các điều kiện quy định tại Điều 193 của Luật Đất đai 2013. Cần tham khảo thêm các văn bản hướng dẫn thi hành để có thông tin đầy đủ và chính xác nhất.

 Hữu ích

 Không hữu ích

c) Ngữ cảnh truy xuất cho chức năng Hỏi-Đáp.

Tương tự chức năng So sánh, đây là bằng chứng cốt lõi cho sự đáng tin cậy của kiến trúc RAG. Người dùng có toàn quyền kiểm tra và đối chiếu nội dung gốc của các điều luật đã được hệ thống sử dụng để tạo ra câu trả lời. Tính năng "Xem các điều luật liên quan nhất đã được sử dụng" không chỉ giúp tăng độ tin cậy mà còn là một công cụ học tập hữu ích, cho phép người dùng tự khám phá sâu hơn các quy định pháp luật liên quan đến vấn đề họ quan tâm.

 Xem các điều luật liên quan nhất đã được sử dụng

Điều 177.0 Luật Đất đai 2024 (Điểm liên quan: 8.0344)

Điều 177.

Hạn mức nhận chuyển quyền sử dụng đất nông nghiệp của cá nhân

1. Hạn mức nhận chuyển quyền sử dụng đất nông nghiệp của cá nhân không quá 15 lần hạn mức giao đất nông nghiệp của cá nhân đối với mỗi loại đất quy định tại các khoản 1, 2 và 3 Điều 176 của Luật này.

2. Việc xác định hạn mức nhận chuyển quyền sử dụng đất nông nghiệp quy định tại khoản 1 Điều này phải căn cứ vào các yếu tố sau đây:

a) Điều kiện về đất đai và công nghệ sản xuất;

b) Chuyển dịch lực lượng lao động, cơ cấu kinh tế, quá trình đô thị hóa.

3. Ủy ban nhân dân cấp tỉnh, căn cứ quy định tại khoản 1 và khoản 2 Điều này, quy định hạn mức nhận chuyển quyền sử dụng đất nông nghiệp của cá nhân phù hợp với điều kiện cụ thể của địa phương.

Điều 130.0 Luật Đất đai 2013 (Điểm liên quan: 8.0145)

Điều 130.

Hạn mức nhận chuyển quyền sử dụng đất nông nghiệp của hộ gia đình, cá nhân

1. Hạn mức nhận chuyển quyền sử dụng đất nông nghiệp của hộ gia đình, cá nhân không quá 10 lần hạn mức giao đất nông nghiệp của hộ gia đình, cá nhân đối với mỗi loại đất quy định tại các khoản 1, 2 và 3 Điều 129 của Luật này.

2. Chính phủ quy định hạn mức nhận chuyển quyền sử dụng đất nông nghiệp của hộ gia đình, cá nhân phù hợp với điều kiện cụ thể theo từng vùng và từng thời kỳ.

Điều 47.0 Luật Đất đai 2024 (Điểm liên quan: 6.7922)

Việc hiển thị rõ ràng ngữ cảnh đã qua bước Rerank (kèm theo điểm số liên quan) là một minh chứng quan trọng cho hiệu quả của pipeline truy xuất hai giai đoạn, đồng thời khẳng định sự vượt trội của hệ thống so với các mô hình LLM thuần hoạt động trong "hộp đen".