

TĂNG CƯỜNG KHẢ NĂNG LẬP LUẬN TRÊN TÀI LIỆU HÌNH ẢNH (DOCVQA) BẰNG CÁCH KẾT HỢP DEEPSEEK OCR VÀ TRUY XUẤT THÔNG TIN PHÁP LÝ

Dương Tấn Thành - 250101063

Tóm tắt

- Lớp: CS2205.CH201
- Link Github của nhóm:
<https://github.com/thanhdt942/CS2205.CH201>
- Link YouTube video: <https://youtu.be/xUoHr3rPplg>
- Họ và tên: Dương Tấn Thành
- MSSV: 250101063

Giới thiệu

Khối lượng lớn tài liệu lưu trữ dạng hình ảnh/scan (Hợp đồng, văn bản).

DocVQA truyền thống

- Trích xuất thông tin đơn giản (Ngày tháng, tên,...) khá tốt
- Tuy nhiên với thông tin phức tạp dạng bản thì chưa xử lý tốt.
- Vấn đề ảo giác (Hallucination) khi tư vấn pháp lý do thiếu văn bản luật chính xác làm căn cứ.



DeepSeek-OCR đã mở ra khả năng hiểu cấu trúc tài liệu vượt trội



Kỹ thuật RAG mang lại khả năng kiểm chứng thông tin tin cậy dựa trên dữ liệu luật Việt Nam (Legal Knowledge Base)

CỘNG HÒA XÃ HỘI CHỦ NGHĨA VIỆT NAM
Độc lập – Tự do – Hạnh phúc
—o—
HỢP ĐỒNG MUA BÁN HÀNG HÓA
Số: 05/2023/HĐMB

- Căn cứ Bộ luật Dân sự số 91/2015/QH13 ngày 24/11/2015 và các văn bản pháp luật liên quan;
- Căn cứ Luật Thương mại số 36/2005/QH11 ngày 14/06/2005 và các văn bản pháp luật liên quan;
- Căn cứ vào nhu cầu và khả năng của các bên;

Hôm nay, ngày 03 tháng 09 năm 2023, chúng tôi gồm có:

BÊN BÁN ("BÊN A"): CÔNG TY KẾ TOÁN MINH VIỆT
Địa chỉ : Số 198 Nguyễn Tuấn, Thanh Xuân, Hà Nội.
Mã số thuế : 0108369272
Đại diện : Bà Lê Thị Minh. Chức vụ : Giám đốc.

BÊN MUA ("BÊN B"): CÔNG TY TNHH THƯƠNG MẠI HOÀNG LUYẾT
Địa chỉ : 135 T3, KĐT Đại Kim, Hoàng Mai, Hà Nội.
Mã số thuế : 0107638111.
Đại diện : Bà Nguyễn Thị Thu. Chức vụ : Giám đốc.

Trên cơ sở thỏa thuận, hai bên thống nhất ký kết hợp đồng mua bán hàng hóa với các điều khoản sau đây:

ĐIỀU 1. TÊN HÀNG - SỐ LƯỢNG - CHẤT LƯỢNG - GIÁ TRỊ HỢP ĐỒNG

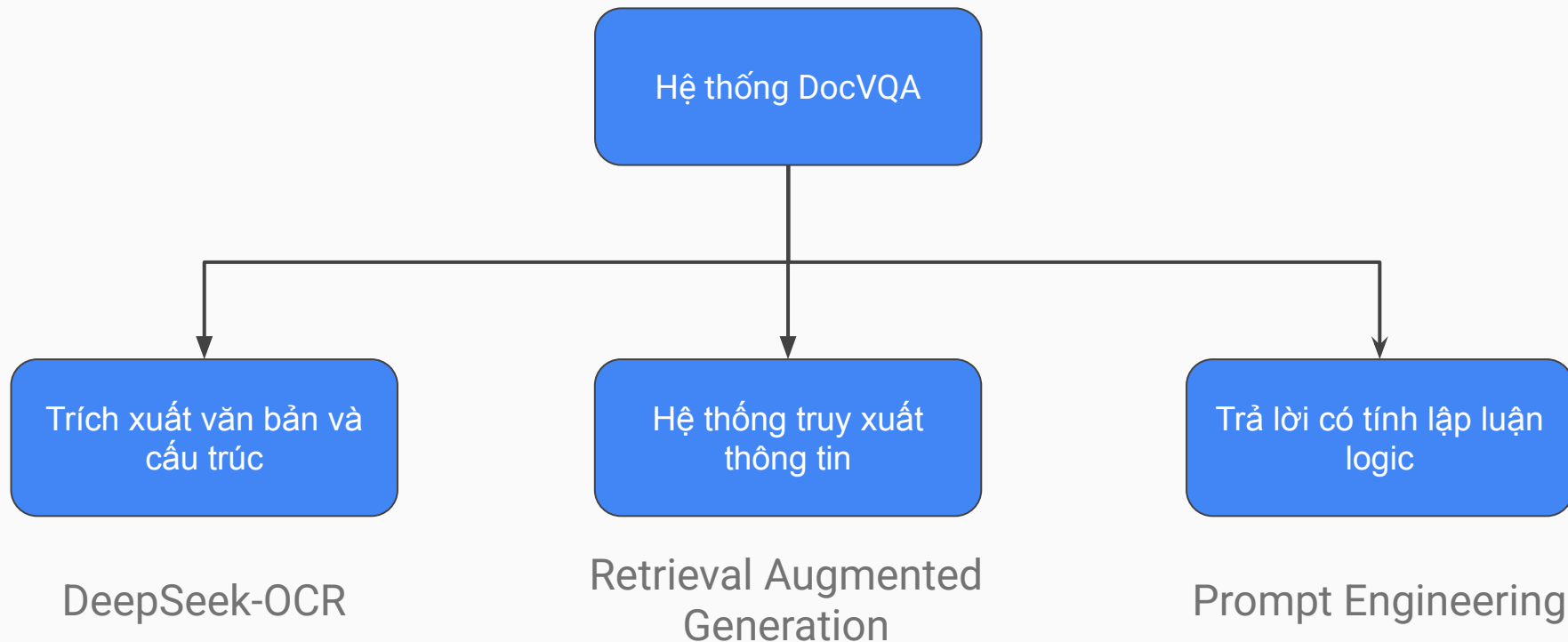
Đơn vị tính: Việt Nam đồng					
STT	Tên hàng hóa	Đơn vị tính	Số lượng	Đơn giá	Thành tiền
01	Ghế xoay X.T18	Chiếc	20	500.000	10.000.000
Cộng giá trị tiền hàng					10.000.000
Thuế GTGT 10%					1.000.000
Tổng giá trị Hợp đồng					11.000.000

Số tiền bằng chữ: Mười một triệu đồng.

ĐIỀU 2. THANH TOÁN.
1. Bên Mua phải thanh toán cho bên Bán số tiền ghi tại Điều 1 của Hợp đồng vào ngày 04/09/2023.
2. Bên Mua thanh toán cho Bên Bán theo hình thức chuyển khoản vào tài khoản của Bên Bán chi tiết như sau:
Số tài khoản: 1000005666.
Nơi tại: Ngân hàng TMCP Thương Tín. Chi nhánh Hoàng Mai.

ĐIỀU 3: THỜI GIAN, ĐỊA ĐIỂM VÀ PHƯƠNG THỨC GIAO HÀNG
1. Việc bán giao ứng đồng cho Bên Mua được thực hiện vào ngày 04/09/2023.
2. Địa điểm giao nhận hàng tại Trụ sở của Bên Mua: Số 198 Nguyễn Tuấn, Thanh Xuân, Hà Nội.

Mục tiêu



Nội dung và Phương pháp

Phân tích thực trạng: Đánh giá hạn chế của OCR truyền thống trên văn bản tiếng Việt phức tạp.

Triển khai DeepSeek-OCR: Ứng dụng mô hình mới cho tác vụ OCR ngữ nghĩa (Semantic OCR).

Xây dựng Legal Knowledge Base:
Tạo Vector Database cho các Bộ luật

Thiết kế Module: Kết hợp thông tin tài liệu và luật để suy luận câu trả lời.

Đề tài tập trung giải quyết 4 vấn đề cốt lõi nhằm xây dựng pipeline xử lý pháp lý

Nội dung và Phương pháp

Giai đoạn 1: Xử lý Đầu vào (Input Processing)

- *Input:* Ảnh tài liệu (Hợp đồng, v.v.).
- *Core Tech:* **DeepSeek-OCR**.
- *Highlight:* Xuất ra định dạng **Markdown** (Giữ nguyên bảng biểu/cấu trúc) => Tạo Document Embedding.

Giai đoạn 2: Truy xuất Tri thức (Legal RAG)

- *Input:* Câu hỏi người dùng (User Query).
- *Core Tech:* **Semantic Search + HyDE**.
- *Process:* Tìm kiếm trong Vector DB Luật Việt Nam.
- *Output:* Top-K điều luật liên quan nhất.

Giai đoạn 3: Suy luận & Sinh câu trả lời (Reasoning)

- *Input:* [Nội dung OCR] + [Top-K Điều luật] + [Câu hỏi].
- *Core Tech:* **LLM + Chain-of-Thought (CoT)**.
- *Action:* Đối chiếu điều khoản tài liệu với quy định pháp luật.
- *Output:* Câu trả lời tư vấn pháp lý cuối cùng.

Kết quả dự kiến

KẾT QUẢ MONG ĐỢI

- Xây dựng thành công hệ thống DocVQA có khả năng đọc hiểu hợp đồng/văn bản hành chính tiếng Việt.
- Hệ thống có khả năng phát hiện các điểm mâu thuẫn giữa văn bản trong ảnh và quy định pháp luật (nếu có).
- Giao diện demo cho phép người dùng upload ảnh, hỏi đáp và xem được trích dẫn luật cụ thể mà hệ thống sử dụng để trả lời.

Tài liệu tham khảo

- [1] [Haoran Wei, Yaofeng Sun, Yukun Li: DeepSeek-OCR: Contexts Optical Compression. CoRR abs/2510.18234 \(2025\)](#)
- [2] [P. Lewis et al., "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks," in NeurIPS, 2020.](#)
- [3] [M. Mathew, D. Karatzas, and C. V. Jawahar, "DocVQA: A Dataset for VQA on Document Images," in WACV, 2021, pp. 2200-2209.](#)
- [4] [Y. Huang et al., "LayoutLMv3: Pre-training for Document AI with Unified Text and Image Masking," in ACM Multimedia, 2022, pp. 4083-4091.](#)
- [5] [K. Nguyen et al., "VLSP 2023 - LTER: A Summary of the Challenge on Legal Textual Entailment Recognition" in VLSP, 2023.](#)