

ỨNG DỤNG THỊ GIÁC MÁY TÍNH VÀ MÔ HÌNH NGÔN NGỮ LỚN ĐỂ TỰ ĐỘNG HOÁ THÔNG MINH VIỆC PHÂN TÍCH VĂN BẢN HÀNH CHÍNH

Mai Chấn Tuấn - 240101030



Tóm tắt

- Link Github của nhóm: <https://github.com/tuanmc15/ppnckh-cuoiki>
- Link YouTube video:
<https://www.youtube.com/watch?v=tY-AplvycBs>
- Ảnh + Họ và Tên của các thành viên

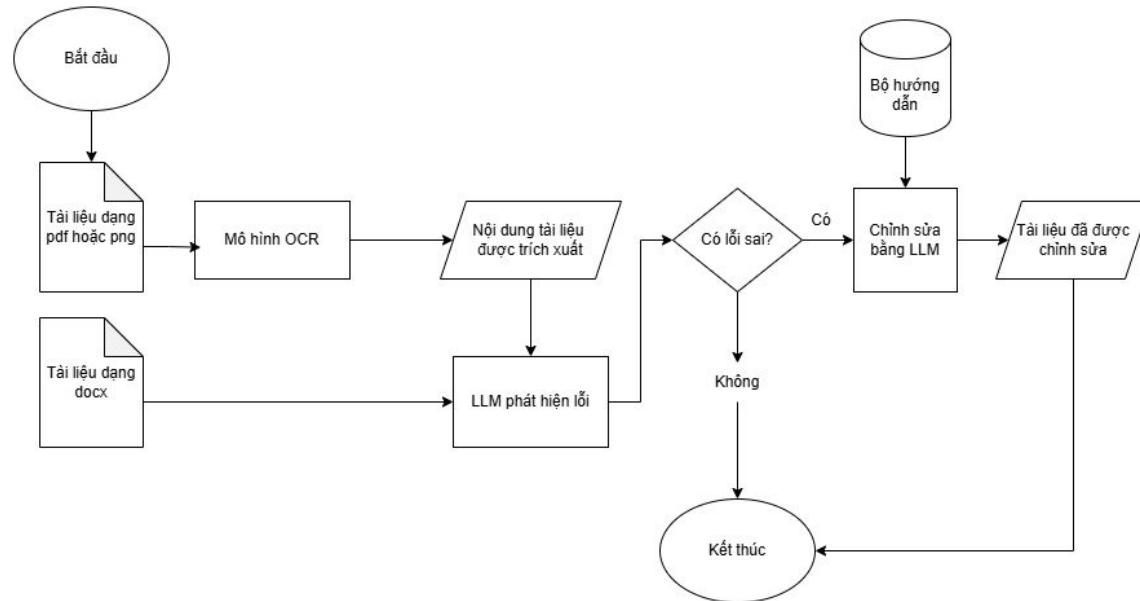
Mai Chấn Tuấn
240101030



Giới thiệu

- Tài liệu hành chính là một trong những loại tài liệu được quản lý chặt chẽ nhất trong nhiều lĩnh vực do vai trò quan trọng của chúng trong việc quản trị và vận hành.
- Những sai sót trong tài liệu hành chính có thể gây ra hậu quả nghiêm trọng, bao gồm tranh chấp pháp lý, quản lý tài chính sai lệch, bị phạt do vi phạm quy định và tổn hại danh tiếng.
- Chúng tôi đề xuất một giải pháp tích hợp công nghệ LLM và thị giác máy tính tiên tiến để phát hiện lỗi trong tài liệu hành chính nhằm ngăn chặn những sai sót đã đề cập

Giới thiệu



Mục tiêu

- **Phân tích và xử lý nội dung tài liệu:** Ứng dụng Optical Character Recognition (OCR) và mô hình ngôn ngữ lớn (LLM) để đọc văn bản, phân tích cú pháp, ngữ pháp và ngữ nghĩa với độ chính xác cao.
- **Đánh giá hiệu suất:** Sử dụng các chỉ số precision, recall, F1 score để đo lường hiệu quả của hệ thống.
- **Triển khai hệ thống:** Sử dụng dịch vụ đám mây như AWS để triển khai Virtual Private Computer nhằm host 1 website cho phép người dùng tương tác với hệ thống một cách thuận tiện.

Nội dung và Phương pháp

- **Nội dung 1: Nghiên cứu cách thức OCR hoạt động.**
 - **Phương pháp:** Nghiên cứu các tài liệu hiện có về công nghệ OCR.
- **Nội dung 2: Sử dụng LLM để phân tích và sửa lỗi văn bản.**
 - **Phương pháp:** Áp dụng mô hình ngôn ngữ lớn (LLM) để phân tích văn bản hành chính dựa trên một bộ hướng dẫn (guideline) được xây dựng nhằm chuẩn hóa các tiêu chí về hình thức và nội dung.
- **Nội dung 3: Triển khai hệ thống trên dịch vụ đám mây**
 - **Phương pháp:** Sử dụng dịch vụ điện toán đám mây của AWS để triển khai hệ thống
- **Nội dung 4: Đánh giá hệ thống.**
 - **Phương pháp:** Sử dụng các chỉ số như precision, recall, accuracy, F1 score, false positive rate, false negative rate,...

Kết quả dự kiến

- Nắm được phần cốt lõi và có thể áp dụng OCR vào hệ thống một cách thuần thục.
- Kết hợp được OCR và LLM nhằm tạo ra một ứng dụng có thể phân tích văn bản hành chính
- Ứng dụng có thể sử dụng trong thực tế, giúp rút ngắn thời gian soạn thảo, giảm thiểu lỗi xuống mức tối đa.
- Ứng dụng có thể được ứng dụng linh hoạt tùy vào bộ hướng dẫn cụ thể của từng tổ chức.

Tài liệu tham khảo

- [1] R. Liu and N. B. Shah, “ReviewerGPT? An Exploratory Study on Using Large Language Models for Paper Reviewing,” arXiv.org, 2023. <https://arxiv.org/abs/2306.00622>
- [2] I. Beltagy, M. E. Peters, and A. Cohan, “Longformer: The Long-Document Transformer,” *arXiv:2004.05150 [cs]*, Dec. 2020, Available: <https://arxiv.org/abs/2004.05150>
- [3] G. Kim et al., “OCR-Free Document Understanding Transformer,” *Lecture Notes in Computer Science*, pp. 498–517, Jan. 2022, doi: https://doi.org/10.1007/978-3-031-19815-1_29.
- [4] J. Wu, Z. Wu, R. Li, A. Hasan, Y. Kim, J. P. Y. Cheung, T. Zhang, and H. Wu, “Integrating Knowledge Retrieval and Large Language Models for Clinical Report Correction,” *Arxiv.org*, 2023. <https://arxiv.org/html/2406.15045v2> (accessed Feb. 27, 2025).
- [5] C. Amrhein and S. Clematide, “Supervised OCR Error Detection and Correction Using Statistical and Neural Machine Translation Methods,” *Journal for Language Technology and Computational Linguistics (JLCL)*, vol. 33, no. 1, pp. 49–76, 2018, doi: <https://www.zora.uzh.ch/id/eprint/162394/1/AmrheinClematide2018.pdf>.
- [6] J. Evershed and K. Fitch, “Correcting noisy OCR,” May 2014, doi: <https://doi.org/10.1145/2595188.2595200>.