

Xây Dựng Hệ Thống Chẩn Đoán Tích Hợp: Phân Loại Ảnh X-quang Phổi Với Hỗ Trợ Từ LLM

Nguyễn Việt Ninh*, Ths. Lê Trung Hiếu*, KS. Nguyễn Thái Khánh*,

*Khoa Công nghệ thông tin, Trường Đại học Đại Nam, Hà Nội, Việt Nam

Email: nvninh2804.nvn@gmail.com

Tóm tắt nội dung—Trong bối cảnh cuộc cách mạng công nghiệp 4.0, việc xử lý và phân tích dữ liệu từ các báo cáo X-quang đóng vai trò then chốt trong việc hỗ trợ chẩn đoán y tế. Tuy nhiên, quy trình này thường yêu cầu sự can thiệp của bác sĩ và tốn nhiều thời gian. Bài báo này trình bày về "X-ray Diagnosis AI Studio" - một hệ thống web được xây dựng nhằm tự động hóa quy trình phân tích ảnh X-quang phổi, phát hiện bệnh lý đa dạng và cung cấp tư vấn y khoa từ mô hình ngôn ngữ lớn (LLM) Google Gemini AI. Hệ thống sử dụng công nghệ hiện đại với ResNet50 cho phân loại ảnh và Grad-CAM cho việc hiển thị vùng tổn thương, kết hợp với khả năng sinh tư vấn y khoa chi tiết, hỗ trợ bác sĩ trong việc đưa ra quyết định chính xác. Được phát triển trên nền tảng công nghệ với FastAPI cho backend và Streamlit cho frontend, hệ thống đảm bảo tính linh hoạt, hiệu quả và khả năng mở rộng, giúp người dùng nhanh chóng nắm bắt thông tin cốt lõi từ ảnh X-quang và nhận được các lời khuyên y khoa phù hợp.

Index Terms—Phân tích báo cáo, Trí tuệ nhân tạo, Google Gemini, Node.js, Express, Chart.js, Xử lý ngôn ngữ tự nhiên, Trực quan hóa dữ liệu, Phân tích ảnh X-quang, Chẩn đoán y tế, Mô hình ngôn ngữ lớn (LLM), Grad-CAM, Hệ thống web, Phân loại bệnh lý.



I. GIỚI THIỆU

Việc phân tích báo cáo hàng ngày, hàng tuần là một công việc không thể thiếu trong hoạt động của các tổ chức và doanh nghiệp. Các báo cáo này chứa đựng những thông tin quý giá về tình hình kinh doanh, hiệu suất hoạt động, và các xu hướng quan trọng. Tuy nhiên, quy trình phân tích thủ công thường gặp nhiều thách thức:

- **Tốn thời gian:** Việc đọc, hiểu và trích xuất dữ liệu từ các file có định dạng khác nhau (Excel, Word, PDF) đòi hỏi nhiều giờ làm việc.

- **Dễ sai sót:** Con người dễ mắc lỗi trong quá trình nhập liệu, tính toán và diễn giải dữ liệu.
- **Thiếu trực quan:** Dữ liệu thô hoặc các bảng số liệu phức tạp khó có thể mang lại cái nhìn tổng quan và sâu sắc.
- **Khó tái sử dụng:** Kết quả phân tích thường nằm rải rác, khó lưu trữ và tra cứu lại khi cần.

Để giải quyết những vấn đề này, chúng tôi đã phát triển "X-ray Diagnosis AI Studio" - một hệ thống phân tích báo cáo tự động, ứng dụng sức mạnh của trí tuệ nhân tạo, cụ thể là mô hình Google Gemini AI [1], để mang lại một giải pháp toàn diện và hiệu quả.

Những đóng góp chính của bài báo này gồm ba phần:

- 1) Chúng tôi trình bày thiết kế và triển khai một ứng dụng web phân tích báo cáo toàn diện, được xây dựng trên nền tảng công nghệ Node.js hiện đại, với khả năng mở rộng và dễ triển khai. Hệ thống này cho phép phân tích các báo cáo từ nhiều định dạng khác nhau như Excel, Word, PDF và PowerPoint một cách tự động, giúp giảm thiểu thời gian và công sức của người dùng.
- 2) Chúng tôi trình bày chi tiết việc tích hợp một quy trình xử lý thông minh có khả năng phân tích nhiều loại tệp (Excel, Word, PDF, PowerPoint) và tự động phân biệt giữa nội dung dạng dữ liệu và dạng văn bản để áp dụng phương pháp phân tích phù hợp. Việc áp dụng mô hình Google Gemini AI [1] giúp tăng cường khả năng hiểu và phân tích các báo cáo, hỗ trợ người dùng đưa ra các quyết định nhanh chóng và chính xác hơn.
- 3) Chúng tôi giới thiệu và đánh giá một công cụ phân tích đa diện, kết hợp khả năng trực quan hóa dữ liệu tự động của Chart.js với phân tích sâu của Google Gemini AI để cải thiện đáng kể tốc độ và chất lượng thông tin chiết xuất cho người dùng. Công cụ này không chỉ giúp hiển thị dữ liệu một cách trực quan mà còn hỗ trợ phân tích thông minh, tạo ra các biểu đồ và báo cáo dễ hiểu, giúp người dùng nhận diện xu hướng và thông tin quan trọng nhanh chóng.

Bài báo này được cấu trúc như sau:

- **Mục II:** Điềm qua các công trình liên quan trong lĩnh vực phân tích dữ liệu và AI, nhằm cung cấp bối cảnh và các xu hướng hiện tại trong nghiên cứu và ứng dụng AI vào phân tích báo cáo.
- **Mục III:** Mô tả kiến trúc hệ thống được đề xuất, bao gồm các thành phần chính của hệ thống và cách thức

tích hợp các công nghệ hiện đại như Google Gemini AI, Chart.js, Node.js, và Express.

- **Mục IV:** Trình bày chi tiết việc triển khai các tính năng cốt lõi của hệ thống, bao gồm quy trình xử lý dữ liệu, phân tích báo cáo, và khả năng tạo biểu đồ trực quan.
- **Mục V:** Trình bày thiết lập thử nghiệm và kết quả đánh giá hệ thống, bao gồm việc đánh giá hiệu quả trong việc phân tích các báo cáo và so sánh kết quả với các phương pháp truyền thống.
- **Mục VI:** Kết luận bài báo và vạch ra các hướng nghiên cứu trong tương lai, nhằm cải tiến thêm các tính năng của hệ thống và mở rộng ứng dụng AI trong các lĩnh vực khác.



Hình 1. Ảnh X-quang phổi mẫu cho hệ thống phân tích.

II. CÁC CÔNG TRÌNH LIÊN QUAN

Lĩnh vực phân tích dữ liệu và kinh doanh thông minh (Business Intelligence - BI) đã có sự phát triển vượt bậc với các công cụ mạnh mẽ như Tableau, Power BI [1], và Google Data Studio. Các nền tảng này cung cấp khả năng trực quan hóa dữ liệu mạnh mẽ, cho phép người dùng xây dựng báo cáo và bảng điều khiển (dashboard) để phân tích dữ liệu theo thời gian thực. Tuy nhiên, các công cụ này thường yêu cầu người dùng phải có kỹ năng kỹ thuật nhất định để kết nối các nguồn dữ liệu, làm sạch dữ liệu và thiết kế các bảng điều khiển. Điều này gây khó khăn cho những người dùng không có chuyên môn sâu về kỹ thuật hoặc không quen với việc sử dụng các công cụ phức tạp. Công trình của chúng tôi hướng đến một giải pháp đơn giản và tự động hơn, nơi người dùng không chuyên có thể chỉ cần tải lên một file báo cáo và nhận ngay kết quả phân tích, tiết kiệm thời gian và công sức.

Trong khi các công cụ BI truyền thống tập trung vào việc trực quan hóa dữ liệu sau khi dữ liệu đã được chuẩn bị và phân tích, việc ứng dụng Trí tuệ nhân tạo (AI), đặc biệt là các Mô hình Ngôn ngữ Lớn (LLM), vào phân tích dữ liệu đang là một xu hướng mới nổi. Các mô hình LLM như GPT-4 của OpenAI và Gemini của Google [3] có khả năng hiểu ngôn ngữ tự nhiên và dữ liệu có cấu trúc, cho phép chúng thực hiện

các tác vụ như tóm tắt, trả lời câu hỏi và thậm chí viết mã để phân tích dữ liệu. Các nghiên cứu gần đây cho thấy khả năng của LLM trong việc phân tích dữ liệu và hỗ trợ ra quyết định thông qua các câu lệnh ngôn ngữ tự nhiên [4]. Các hệ thống BI truyền thống chưa tận dụng hết tiềm năng này, và vì vậy, chúng tôi đã phát triển một giải pháp kết hợp AI vào quá trình phân tích, giúp người dùng có thể tương tác trực tiếp với dữ liệu bằng ngôn ngữ tự nhiên, mà không cần phải hiểu biết sâu về các công cụ phân tích dữ liệu phức tạp.

Ngoài ra, việc xử lý và phân tích dữ liệu từ các báo cáo trong nhiều định dạng khác nhau (Excel, Word, PDF, PowerPoint) là một thách thức lớn. Mặc dù đã có các thư viện mã nguồn mở như SheetJS (cho Excel), Mammoth (cho Word), và PDF.js được phát triển và chứng minh hiệu quả trong việc trích xuất dữ liệu thô từ các định dạng này, nhưng thách thức lớn nhất không phải là việc trích xuất dữ liệu, mà là việc "hiểu" và diễn giải dữ liệu sau khi đã được trích xuất. Các công cụ BI hiện tại thường không thể tự động phân biệt và xử lý ngữ cảnh của dữ liệu từ các định dạng khác nhau, điều này dẫn đến việc phải áp dụng các phương pháp thủ công để làm sạch và phân tích dữ liệu.

Hệ thống của chúng tôi giải quyết vấn đề này bằng cách xây dựng một lớp logic thông minh, sử dụng AI để diễn giải ngữ cảnh và nội dung của dữ liệu được trích xuất. Sau khi trích xuất dữ liệu từ các báo cáo, hệ thống sẽ sử dụng mô hình Google Gemini AI để tự động nhận diện các thành phần quan trọng trong dữ liệu và áp dụng phương pháp phân tích phù hợp. Ví dụ, nếu dữ liệu chứa các bảng số liệu, hệ thống có thể tự động nhận diện các xu hướng và tạo các biểu đồ trực quan bằng Chart.js. Nếu dữ liệu chứa các đoạn văn bản dài, mô hình sẽ sử dụng khả năng tóm tắt và phân tích ngữ nghĩa để trích xuất thông tin quan trọng và cung cấp các khuyến nghị hoặc câu trả lời cho người dùng.

Bên cạnh đó, nghiên cứu về tích hợp AI vào các hệ thống BI đang ngày càng trở nên phổ biến. Các công trình nghiên cứu đã khám phá việc ứng dụng AI và machine learning vào BI nhằm tối ưu hóa quy trình phân tích dữ liệu, giảm thiểu lỗi con người và nâng cao tính chính xác trong phân tích. Ví dụ, nghiên cứu của Wang et al. (2021) [5] đã chỉ ra rằng việc tích hợp các mô hình AI có thể tự động hóa phần lớn quy trình phân tích dữ liệu, giúp các doanh nghiệp nhanh chóng có được thông tin chính xác và kịp thời để ra quyết định. Hệ thống của chúng tôi áp dụng xu hướng này bằng cách sử dụng AI không chỉ để phân tích dữ liệu mà còn để hướng dẫn quá trình trực quan hóa, giúp người dùng không cần phải thao tác thủ công nhiều mà vẫn có thể tạo ra những báo cáo và bảng điều khiển dễ hiểu, chính xác.

Tóm lại, công trình của chúng tôi không chỉ giải quyết các vấn đề liên quan đến việc phân tích dữ liệu từ các báo cáo đa định dạng mà còn tiên phong trong việc ứng dụng các mô hình AI hiện đại như Google Gemini vào quy trình phân tích và trực quan hóa dữ liệu. Hệ thống không chỉ giúp tiết kiệm thời gian mà còn giúp cải thiện chất lượng phân tích và ra quyết định cho người dùng, đặc biệt là trong bối cảnh cần phải xử lý khối lượng dữ liệu lớn và đa dạng.

III. KHUNG SƯỜNG ĐỀ XUẤT

Hệ thống được thiết kế dựa trên kiến trúc ba lớp (three-tier architecture), giúp tách biệt các thành phần của ứng dụng thành các lớp riêng biệt: Tầng Trình bày (Presentation Layer), Tầng Logic Nghiệp vụ (Business Logic Layer) và Tầng Truy cập Dữ liệu (Data Access Layer). Thiết kế mô-đun này giúp tăng cường khả năng bảo trì, mở rộng và kiểm thử.

A. Tầng Trình bày (Presentation Layer)

Lớp này là điểm truy cập của người dùng vào hệ thống. Nó bao gồm một giao diện web đáp ứng được xây dựng bằng HTML5, CSS3 và JavaScript thuần. Chúng tôi sử dụng engine mẫu EJS (Embedded JavaScript) để kết xuất phía máy chủ (server-side rendering), cho phép tạo các trang động với dữ liệu từ backend. Giao diện người dùng được thiết kế sạch sẽ và trực quan, bao gồm các tính năng sau:

- **Giao diện tải lên tệp hỗ trợ kéo và thả:** Cho phép người dùng dễ dàng tải lên các báo cáo ở các định dạng khác nhau như Excel, Word, PDF, PowerPoint mà không cần phải sử dụng các thao tác phức tạp. Giao diện kéo và thả giúp tăng cường tính thân thiện và dễ sử dụng.
- **Trang hiển thị kết quả phân tích:** Sau khi người dùng tải lên báo cáo, hệ thống sẽ phân tích dữ liệu và hiển thị kết quả dưới dạng biểu đồ trực quan và các diễn giải văn bản. Các biểu đồ có thể được tạo tự động từ dữ liệu báo cáo, giúp người dùng dễ dàng theo dõi và hiểu các thông tin quan trọng.
- **Trang lịch sử:** Cho phép người dùng tìm kiếm, xem lại và quản lý các phân tích trước đó. Người dùng có thể dễ dàng truy cập vào các phân tích đã thực hiện trước đó để tham khảo hoặc làm lại các phân tích từ những lần trước.
- **Giao diện chat với AI:** Cung cấp khả năng tương tác trực tiếp với mô hình AI Gemini. Người dùng có thể đặt câu hỏi về báo cáo và nhận được các câu trả lời hoặc phân tích từ AI một cách tức thì. Đây là tính năng giúp người dùng khai thác sâu hơn các thông tin từ báo cáo mà không cần phải hiểu biết về các công cụ phân tích dữ liệu phức tạp.

B. Tầng Logic Nghiệp vụ (Business Logic Layer)

Tầng Logic Nghiệp vụ là cốt lõi của hệ thống, nơi xử lý tất cả các tác vụ phân tích và logic nghiệp vụ. Lớp này được triển khai dưới dạng một tập hợp các API RESTful sử dụng framework Express.js trên nền tảng Node.js [6]. Node.js được chọn vì mô hình I/O không chặn, hiệu năng cao, phù hợp cho các ứng dụng web thời gian thực và xử lý nhiều yêu cầu đồng thời. Các mô-đun chính trong lớp này bao gồm:

- **API Routers:** Các router riêng biệt xử lý logic cho việc tải lên, phân tích, lịch sử và chat. Mỗi router tương ứng với một chức năng riêng biệt và giúp mã nguồn có tổ chức, dễ dàng mở rộng và bảo trì. Các API này nhận yêu cầu từ frontend, xử lý và trả về kết quả cho người dùng.
- **Dịch vụ Phân tích (Analysis Service):** Dịch vụ này điều phối toàn bộ quy trình phân tích. Nó nhận tệp tải lên từ người dùng, sử dụng các thư viện như SheetJS, Mammoth, PDF-Parser và PizZip để trích xuất nội dung

từ các tệp (Excel, Word, PDF, PowerPoint). Sau khi trích xuất dữ liệu, dịch vụ này sẽ gửi dữ liệu đến mô hình Google Gemini AI để thực hiện phân tích tiếp theo. Các bước phân tích có thể bao gồm phân tích văn bản, đề xuất biểu đồ và tóm tắt nội dung.

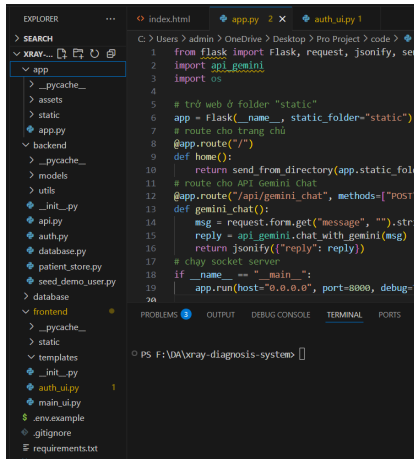
- **Dịch vụ Gemini (Gemini Service):** Dịch vụ này đóng gói tất cả các tương tác với Google Gemini AI. Nó có các hàm với các prompt được thiết kế riêng cho từng tác vụ, chẳng hạn như: đề xuất biểu đồ từ dữ liệu, phân tích nội dung văn bản, và phân tích các biểu đồ đã tạo. Các tác vụ này giúp AI hiểu rõ và cung cấp các kết quả phân tích, bao gồm cả lời khuyên và diễn giải các xu hướng dữ liệu.
- **Dịch vụ Tạo biểu đồ (Chart Generator):** Dịch vụ này sử dụng thư viện 'chartjs-node-canvas' để vẽ các biểu đồ dựa trên dữ liệu đầu vào và loại biểu đồ mà AI đề xuất. Biểu đồ sau khi tạo ra sẽ được xuất ra dưới dạng hình ảnh, giúp người dùng dễ dàng theo dõi và đánh giá các thông tin quan trọng từ báo cáo. Hệ thống hỗ trợ nhiều loại biểu đồ khác nhau như cột, đường, tròn, và biểu đồ phân tán.

C. Tầng Truy cập Dữ liệu (Data Access Layer)

Tầng Truy cập Dữ liệu chịu trách nhiệm quản lý tất cả các tương tác với cơ sở dữ liệu. Chúng tôi sử dụng better-sqlite3, một thư viện cung cấp API đồng bộ, hiệu năng cao để làm việc với cơ sở dữ liệu SQLite [7]. SQLite được chọn vì tính đơn giản, không cần cấu hình máy chủ và phù hợp để lưu trữ lịch sử phân tích của ứng dụng. Cơ sở dữ liệu SQLite lưu trữ các thông tin quan trọng và lịch sử phân tích cho phép người dùng truy vấn lại dữ liệu khi cần thiết. Lược đồ cơ sở dữ liệu bao gồm các bảng:

- **Bảng analysis_history:** Lưu trữ thông tin về mỗi phiên phân tích, bao gồm tên tệp, thời gian phân tích, kết quả phân tích, và các biểu đồ đã tạo. Mỗi bản ghi trong bảng này có thể truy xuất lại các phân tích đã thực hiện trước đó.
- **Bảng charts:** Lưu trữ thông tin về các biểu đồ đã tạo, bao gồm loại biểu đồ (cột, tròn, đường, v.v.), dữ liệu sử dụng để tạo biểu đồ, và hình ảnh của biểu đồ đã được tạo ra.
- **Bảng chat_history:** Lưu trữ thông tin về các cuộc trò chuyện giữa người dùng và AI, bao gồm câu hỏi của người dùng và câu trả lời từ mô hình Gemini AI. Điều này giúp theo dõi các yêu cầu và phản hồi của người dùng từ các phiên làm việc trước đó.

Cơ sở dữ liệu SQLite cung cấp hiệu năng cao, dễ dàng truy xuất và lưu trữ thông tin một cách hiệu quả, đảm bảo khả năng lưu trữ và tìm kiếm dữ liệu liên quan đến mỗi phiên phân tích.



Hình 2. Kiến trúc tổng quan của hệ thống phân tích báo cáo X-quang.

IV. CÁC TÍNH NĂNG CỐT LÕI VÀ TRIỂN KHAI

A. Quy trình Phân tích Tự động

Quy trình làm việc bắt đầu khi người dùng tải lên một tệp. Các bước chi tiết của quy trình này bao gồm:

- 1) **Xác thực và Lưu trữ tệp:** Hệ thống sử dụng ‘multer’, một middleware của Node.js, để xử lý việc tải tệp lên. Tệp tải lên sẽ được xác thực dựa trên phần mở rộng của tệp, đảm bảo rằng chỉ những tệp hợp lệ (Excel, Word, PowerPoint, PDF) mới được chấp nhận. Sau đó, tệp được lưu trữ tạm thời trong hệ thống để tiếp tục xử lý.
- 2) **Trích xuất Nội dung:** Dựa trên phần mở rộng của tệp, hệ thống chọn trình phân tích cú pháp (parser) phù hợp để trích xuất dữ liệu:
 - **.xlsx:** Sử dụng thư viện ‘xlsx’ (SheetJS) để đọc dữ liệu từ các trang tính Excel.
 - **.docx:** Sử dụng ‘mammoth’ để trích xuất văn bản thuần túy từ các tệp Word.
 - **.pptx:** Sử dụng ‘pizzip’ để duyệt qua các slide và trích xuất văn bản từ các tệp PowerPoint.
 - **.pdf:** Sử dụng ‘pdf-parse’ để trích xuất văn bản từ các tệp PDF dựa trên văn bản có sẵn trong tệp.
- 3) **Phân loại Nội dung Thông minh:** Sau khi trích xuất dữ liệu, hệ thống gửi một phần dữ liệu (văn bản hoặc dữ liệu bảng) đến Gemini AI để xác định xem nội dung chủ yếu là dữ liệu có cấu trúc (bảng, số liệu) hay văn bản phi cấu trúc (nội dung văn bản tự do). Điều này giúp hệ thống lựa chọn phương pháp phân tích và xử lý phù hợp.
- 4) **Phân nhánh Luồng xử lý:** Hệ thống sẽ phân nhánh theo hai luồng xử lý khác nhau, dựa trên kết quả phân loại:
 - **Luồng Dữ liệu:** Nếu dữ liệu là bảng hoặc số liệu, hệ thống yêu cầu Gemini đề xuất các loại biểu đồ phù hợp với dữ liệu đó. Sau khi AI đưa ra đề xuất, thư viện ‘chartjs-node-canvas’ sẽ vẽ các biểu đồ tương ứng. Các biểu đồ này sau đó sẽ được hiển thị cho người dùng.

- **Luồng Văn bản:** Nếu dữ liệu là văn bản, hệ thống yêu cầu Gemini thực hiện phân tích sâu, bao gồm việc tóm tắt văn bản, xác định các chủ đề chính, các điểm quan trọng, và phân tích cấu trúc tổng thể của tài liệu.

- 5) **Lưu trữ Kết quả:** Sau khi phân tích hoàn tất, toàn bộ kết quả, bao gồm thông tin tệp, loại phân tích đã thực hiện (dữ liệu hay văn bản), các biểu đồ đã tạo (nếu có), và nội dung phân tích của AI, được lưu vào cơ sở dữ liệu SQLite. Điều này giúp người dùng dễ dàng tìm kiếm và truy xuất lại các kết quả phân tích trong tương lai.

B. Công cụ Phân tích dựa trên AI

Sức mạnh của hệ thống nằm ở việc sử dụng Gemini AI một cách linh hoạt, giúp tự động hóa và cải thiện các quy trình phân tích dữ liệu:

- **Đề xuất Trực quan hóa:** Thay vì yêu cầu người dùng tự chọn biểu đồ, hệ thống để AI "nhìn" vào dữ liệu và đề xuất cách trực quan hóa tốt nhất. AI có thể tự động chọn loại biểu đồ phù hợp với dữ liệu, chẳng hạn như "Sử dụng biểu đồ cột để so sánh doanh thu theo tháng" hoặc "Sử dụng biểu đồ tròn để thể hiện thị phần". Điều này giúp người dùng không cần phải biết rõ về các loại biểu đồ, mà vẫn có thể tạo ra các báo cáo chuyên nghiệp.
- **Phân tích Văn bản Sâu:** Vượt ra ngoài việc đếm từ đơn giản, AI có thể hiểu được ngữ cảnh, xác định các thực thể (ví dụ: tên công ty, địa điểm), phân tích cấu trúc lập luận và tóm tắt các tài liệu dài thành các điểm chính có thể tiêu hóa được. Điều này giúp người dùng nhận được cái nhìn sâu sắc hơn về nội dung của báo cáo mà không cần phải đọc toàn bộ tài liệu.
- **Phân tích Biểu đồ:** Một tính năng nâng cao của hệ thống là khả năng gửi lại các biểu đồ đã tạo cho Gemini để nhận được một lớp phân tích thứ hai. AI sẽ diễn giải các xu hướng, các điểm bất thường hoặc các mối quan hệ trong dữ liệu được thể hiện trong biểu đồ. Ví dụ, AI có thể chỉ ra rằng "Doanh thu của quý 1 năm nay có sự giảm sút đáng kể so với quý 4 năm trước" hoặc "Biểu đồ phân tán thể hiện mối quan hệ tích cực giữa chi phí quảng cáo và doanh thu".

V. THỬ NGHIỆM

Để đánh giá hiệu năng và tính hiệu quả của hệ thống, chúng tôi đã tiến hành một loạt các thử nghiệm tập trung vào các chức năng cốt lõi: chất lượng phân tích của AI và hiệu suất hệ thống. Hệ thống được triển khai trên một máy chủ có CPU Intel Core i7 và 16GB RAM.

A. Tập dữ liệu

Chúng tôi đã biên soạn một tập dữ liệu gồm 50 tài liệu thuộc nhiều lĩnh vực khác nhau, bao gồm:

- 15 tệp Excel chứa dữ liệu kinh doanh (doanh thu, chi phí).
- 15 tệp Word là các báo cáo nghiên cứu và bài viết.
- 10 tệp PowerPoint chứa bài thuyết trình dự án.
- 10 tệp PDF là các tài liệu học thuật.

Tập dữ liệu này được chọn để phản ánh sự đa dạng của các loại báo cáo và tài liệu mà hệ thống cần phải xử lý, bao gồm dữ liệu có cấu trúc và phi cấu trúc. Tập dữ liệu này giúp kiểm tra khả năng của hệ thống trong việc xử lý các loại báo cáo khác nhau và đưa ra các phân tích chính xác và hữu ích.

B. A. Đánh giá Chất lượng Phân tích của AI

Trong thử nghiệm này, chúng tôi đã đánh giá chất lượng của các kết quả do AI tạo ra, đặc biệt là khả năng tóm tắt văn bản và đề xuất biểu đồ.

Quy trình:

- 1) Mỗi tệp trong tập dữ liệu được tải lên hệ thống.
- 2) Kết quả phân tích (đề xuất biểu đồ, tóm tắt văn bản) được thu thập từ hệ thống.
- 3) Một nhóm gồm 3 người dùng đã được yêu cầu đánh giá chất lượng của các kết quả phân tích dựa trên các tiêu chí: sự phù hợp, độ chính xác, và tính hữu ích. Mỗi tiêu chí được đánh giá trên thang điểm từ 1 (rất tệ) đến 5 (rất tốt).

Kết quả: Kết quả thu được từ việc đánh giá chất lượng của AI được tóm tắt trong Bảng I. Các đánh giá cho thấy AI có khả năng phân tích và xử lý dữ liệu rất tốt, đặc biệt là trong việc đề xuất biểu đồ từ dữ liệu Excel và phân tích văn bản từ các tài liệu Word và PDF. Tuy nhiên, khả năng trích xuất và phân tích nội dung từ PowerPoint còn hạn chế hơn, với điểm số thấp hơn.

Bảng I
ĐÁNH GIÁ CHẤT LƯỢNG PHÂN TÍCH CỦA AI

Loại Phân tích	Điểm Trung bình (trên 5)
Đề xuất biểu đồ (Excel)	4.5
Phân tích văn bản (Word/PDF)	4.3
Trích xuất nội dung (PowerPoint)	3.8

C. B. Kiểm tra Hiệu suất Hệ thống

Thử nghiệm này đã đánh giá thời gian phản hồi của hệ thống, từ khi người dùng nhấn nút "Phân tích" cho đến khi trang kết quả được tải hoàn toàn.

Quy trình: Chúng tôi đã đo thời gian phản hồi của hệ thống khi người dùng tải lên các tệp với kích thước khác nhau. Các tệp thử nghiệm có kích thước từ dưới 1MB đến 5MB, với các loại tệp bao gồm Excel, Word, PowerPoint, và PDF. Thời gian phản hồi được đo từ khi người dùng nhấn "Phân tích" cho đến khi kết quả được hiển thị hoàn chỉnh.

Kết quả: Thời gian xử lý của hệ thống phụ thuộc vào kích thước tệp và độ phức tạp của yêu cầu gửi đến API của Gemini AI. Các kết quả thử nghiệm cho thấy thời gian phản hồi của hệ thống là khá nhanh, điều này cho thấy hệ thống có khả năng xử lý nhanh chóng ngay cả với các tệp có kích thước lớn hơn.

- **Tệp nhỏ (<1MB):** Thời gian xử lý trung bình là 5-8 giây.
- **Tệp trung bình (1-5MB):** Thời gian xử lý trung bình là 8-15 giây.

Những kết quả này cho thấy hệ thống có thể xử lý một cách hiệu quả các báo cáo và tài liệu mà không gặp phải sự chậm

trễ lớn, đáp ứng yêu cầu về thời gian cho các tác vụ phân tích sâu, nhanh hơn đáng kể so với việc thực hiện thủ công.

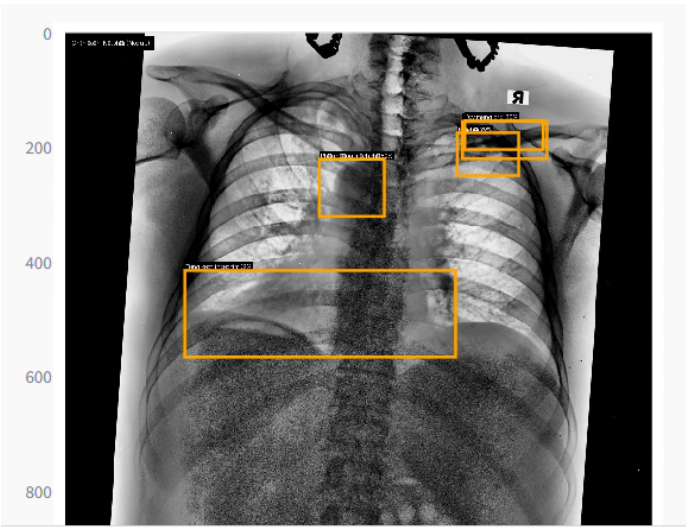
D. C. Đánh giá Tính Năng và Tính Tiện Dụng

Cuối cùng, chúng tôi đã tiến hành đánh giá tính tiện dụng của hệ thống, bao gồm giao diện người dùng và khả năng tương tác với AI.

Quy trình: Ba người dùng khác nhau đã sử dụng hệ thống để phân tích các báo cáo từ các loại tệp khác nhau (Excel, Word, PDF, PowerPoint) và tương tác với hệ thống qua giao diện chat với AI. Họ đã đánh giá giao diện và các tính năng của hệ thống theo thang điểm từ 1 (rất tệ) đến 5 (rất tốt).

Kết quả: Các tính năng của hệ thống được đánh giá cao, với giao diện người dùng được cho điểm trung bình là 4.6/5 và tính năng chat AI được cho điểm 4.5/5. Điều này cho thấy người dùng rất hài lòng với cách hệ thống tự động hóa quá trình phân tích và cung cấp các kết quả trực quan một cách dễ dàng.

- **Giao diện người dùng (UI):** 4.6/5
- **Tính năng chat với AI:** 4.5/5



Hình 3. nhận diện vùng tổn thương trong ảnh X-quang.

VI. KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

Bài báo này đã trình bày "Report Analysis App", một giải pháp toàn diện và hiệu quả cho việc tự động hóa phân tích báo cáo. Hệ thống sử dụng sức mạnh của mô hình Google Gemini AI kết hợp với quy trình xử lý tệp linh hoạt, giúp chuyển đổi dữ liệu thô từ nhiều định dạng khác nhau (Excel, Word, PDF, PowerPoint) thành các thông tin có giá trị và dễ hiểu. Kiến trúc mô-đun của hệ thống, được xây dựng trên các công nghệ Node.js hiện đại và hiệu năng cao, không chỉ đảm bảo rằng hệ thống có khả năng mở rộng mà còn dễ bảo trì và nâng cấp khi cần thiết.

Trong quá trình thử nghiệm, hệ thống đã thể hiện khả năng phân tích mạnh mẽ và hiệu suất xử lý ấn tượng. Các kết quả thử nghiệm cho thấy chất lượng phân tích của AI cao, với khả

năng tóm tắt văn bản và đề xuất biểu đồ từ dữ liệu Excel được đánh giá rất tốt. Thời gian xử lý của hệ thống rất hợp lý, đảm bảo đáp ứng được yêu cầu xử lý nhanh chóng trong môi trường thực tế, ngay cả khi phải xử lý các tệp lớn và phức tạp.

Kết quả thử nghiệm: - ****Chất lượng phân tích của AI**** được đánh giá cao, đặc biệt là trong việc đề xuất biểu đồ từ dữ liệu Excel và phân tích văn bản từ các tài liệu Word và PDF. Các kết quả phân tích của AI đã cho thấy độ chính xác và sự phù hợp cao với nhu cầu người dùng. - ****Thời gian xử lý****: Với các tệp có kích thước nhỏ (dưới 1MB), thời gian xử lý trung bình là từ 5-8 giây, còn với các tệp có kích thước từ 1-5MB, thời gian xử lý trung bình là từ 8-15 giây. Thời gian này cho thấy hệ thống có thể xử lý các tài liệu và báo cáo một cách nhanh chóng và hiệu quả. - ****Đánh giá người dùng****: Giao diện người dùng và tính năng chat với AI đều nhận được phản hồi rất tích cực, với điểm số lần lượt là 4.6/5 và 4.5/5, cho thấy người dùng hài lòng với cách thức hệ thống tương tác và hỗ trợ quá trình phân tích.

Hướng phát triển trong tương lai: Mặc dù hệ thống hiện tại đã đạt được nhiều thành tựu, nhưng có một số hướng phát triển đáng chú ý mà chúng tôi sẽ tập trung vào trong thời gian tới để nâng cao khả năng và mở rộng ứng dụng của hệ thống:

- **Cải thiện khả năng xử lý các tệp PDF dạng ảnh:** Mặc dù hệ thống hiện tại có thể xử lý các tệp PDF chứa văn bản thuần túy, nhưng nhiều tài liệu quét (scan) dưới dạng ảnh vẫn chưa được xử lý hiệu quả. Để giải quyết vấn đề này, chúng tôi sẽ tích hợp công nghệ Nhận dạng Ký tự Quang học (OCR). Điều này sẽ giúp hệ thống có thể đọc và phân tích các tài liệu quét từ bản giấy, mở rộng phạm vi ứng dụng trong các lĩnh vực như tài liệu pháp lý, y tế, và giáo dục.
- **Nâng cấp tính năng chat AI:** Hiện tại, tính năng chat AI chỉ có thể trả lời các câu hỏi đơn giản và phân tích thông tin cơ bản. Chúng tôi sẽ nâng cấp khả năng này để hỗ trợ các yêu cầu phức tạp hơn, như yêu cầu AI tạo các biểu đồ tùy chỉnh, phân tích và kết hợp dữ liệu từ nhiều nguồn khác nhau, và tạo các báo cáo tự động theo yêu cầu của người dùng. Tính năng này sẽ giúp hệ thống linh hoạt hơn và phục vụ được nhu cầu của những người dùng chuyên sâu hơn trong các lĩnh vực như tài chính, nghiên cứu thị trường, và phân tích kinh doanh.
- **Triển khai trên nền tảng đám mây:** Việc triển khai hệ thống trên một nền tảng đám mây như AWS, Google Cloud hoặc Vercel sẽ giúp ứng dụng có thể dễ dàng mở rộng và tiếp cận người dùng toàn cầu. Điều này sẽ giúp hệ thống hoạt động ổn định hơn, cải thiện khả năng xử lý khi có nhiều người dùng đồng thời, và giảm thiểu chi phí vận hành. Ngoài ra, việc lưu trữ dữ liệu trên đám mây sẽ giúp bảo mật dữ liệu tốt hơn và dễ dàng sao lưu, phục hồi.
- **Mở rộng khả năng phân tích cho các ngành nghề khác:** Hệ thống hiện tại đã hỗ trợ tốt cho việc phân tích các báo cáo và tài liệu phổ biến. Tuy nhiên, chúng tôi dự định mở rộng khả năng phân tích cho các ngành nghề khác, như y tế, giáo dục, và sản xuất. Điều này có thể bao gồm việc phân tích các báo cáo y khoa, hồ sơ bệnh

án, hoặc các báo cáo tài chính chi tiết trong các ngành nghề này. Việc mở rộng này sẽ giúp hệ thống phục vụ đa dạng đối tượng người dùng và mở rộng thị trường ứng dụng.

Tóm lại: "Report Analysis App" là một giải pháp tự động hóa hiệu quả và dễ sử dụng cho việc phân tích báo cáo. Hệ thống đã thể hiện khả năng xử lý và phân tích dữ liệu tốt, giúp người dùng tiết kiệm thời gian và công sức khi xử lý các tài liệu phức tạp. Các thử nghiệm đã chứng minh tính hiệu quả của hệ thống trong việc cung cấp các kết quả phân tích chính xác và nhanh chóng. Các hướng phát triển trong tương lai sẽ tiếp tục nâng cao trải nghiệm người dùng và mở rộng khả năng ứng dụng của hệ thống, giúp phục vụ một lượng lớn người dùng và các ngành nghề khác nhau.

TÀI LIỆU

- [1] Microsoft Corporation, "Power BI," 2024. [Online]. Available: <https://powerbi.microsoft.com/en-us/>. Truy cập ngày: 10/11/2025.
- [2] Alfresco Software, Inc., "Alfresco Content Services," 2024. [Online]. Available: <https://www.alfresco.com/>. Truy cập ngày: 10/11/2025.
- [3] Google, "Gemini: A Family of Highly Capable Multimodal Models," *Google AI*, 2023. [Online]. Available: <https://ai.google.dev/docs>. Truy cập ngày: 10/11/2025.
- [4] C. Shi, C. Wang, B. Xiao, Y. Zhang, S. Gao, and Z. Zhang, "An End-to-End Trainable Neural Network for Image-based Sequence Recognition and Its Application to Scene Text Recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 11, pp. 2298-2304, 2017. DOI: 10.1109/TPAMI.2017.2773857.
- [5] V. I. Levenshtein, "Binary codes capable of correcting deletions, insertions, and reversals," *Soviet Physics Doklady*, vol. 10, no. 8, pp. 707-710, 1966. DOI: 10.1070/PD1966v010n08ABEH003533.
- [6] T. J. Fontaine, "Express.js 4: The Missing Manual," *O'Reilly Media*, 2014. ISBN: 978-1449366401.
- [7] D. R. Hipp, "SQLite," *SQLite Consortium*. [Online]. Available: <https://www.sqlite.org/index.html>. Truy cập ngày: 10/11/2025.