

ỨNG DỤNG AI TRONG TỔNG HỢP THÔNG TIN THỊ TRƯỜNG VÀ PHÂN TÍCH CÁC CHỈ SỐ KINH TẾ VĨ MÔ

Đào Hương Giang, Đỗ Hồng Hạnh

Viện Dầu khí Việt Nam (VPI)

Email: giang.dh@vpi.pvn.vn

<https://doi.org/10.47800/PVSI.2025.03-06>

Tóm tắt

Phân tích dữ liệu trong lĩnh vực kinh tế đang chuyển từ các phương pháp truyền thống sang hiện đại dựa trên trí tuệ nhân tạo (AI) giúp tự động hóa quá trình thu thập và xử lý dữ liệu, có khả năng dự báo sự biến động của thị trường, là cơ sở để doanh nghiệp đưa ra quyết định chính xác hơn. Bài viết giới thiệu kết quả ứng dụng AI tại Viện Dầu khí Việt Nam (VPI) để phát triển sản phẩm phân tích số tổng hợp thông tin thị trường, giúp nâng cao hiệu quả công tác tổng hợp, xử lý dữ liệu tự động; dự báo chỉ tiêu kinh tế và cung cấp trợ lý ảo hỗ trợ hỏi đáp thông tin thị trường.

Từ khóa: Kinh tế vĩ mô, trí tuệ nhân tạo, thị trường, xử lý dữ liệu, trợ lý ảo.

1. Giới thiệu

Những thay đổi trong nền kinh tế tác động mạnh đến nhiều lĩnh vực, bao gồm năng lượng nói chung và dầu khí nói riêng. Việc phân tích dữ liệu kinh tế không chỉ cung cấp bức tranh toàn cảnh nhằm hỗ trợ theo dõi, đánh giá diễn biến nền kinh tế mà còn là dữ liệu đầu vào quan trọng phục vụ công tác nghiên cứu về thị trường năng lượng và dầu khí.

Phân tích dữ liệu trong lĩnh vực kinh tế đang chuyển từ các phương pháp truyền thống sang hiện đại dựa trên công nghệ như trí tuệ nhân tạo (AI). Việc tích hợp AI vào phân tích dữ liệu đã tạo ra sự đổi mới mang tính cách mạng trong hoạt động của các tổ chức tài chính, chủ yếu nhờ vào khả năng xử lý, phân tích và diễn giải các bộ dữ liệu khổng lồ [1]. AI đã được sử dụng trong kinh tế học cho các mục đích dự báo, phân tích thị trường và đánh giá tác động của các chính sách [2]; hoạt động tài chính và quy trình phân tích thông tin cho các giao dịch thương mại [3]; hoặc phân tích thông tin tài chính [4]. Xu hướng này khẳng định vai trò ngày càng quan trọng của AI trong việc nâng cao hiệu quả và độ chính xác trong phân tích dữ liệu kinh tế.

Thay vì tra cứu thủ công từ 3 - 5 nguồn báo cáo khác nhau, AI có thể hỗ trợ tìm kiếm thông tin, giúp tiết kiệm

50 - 70% thời gian thực hiện. Đặc biệt, sản phẩm số tổng hợp, xử lý và lưu trữ dữ liệu vào cơ sở dữ liệu có cấu trúc. Sau khi dữ liệu được thu thập, AI hỗ trợ xử lý theo nhu cầu người dùng và lưu trữ vào cơ sở dữ liệu có cấu trúc, giúp dễ dàng truy xuất và phân tích sau này.

Với việc trực quan hóa thông tin trên báo cáo, sản phẩm cho phép người dùng dễ dàng theo dõi các chỉ số kinh tế vĩ mô, được thiết kế với giao diện dễ sử dụng, có thể tùy chỉnh để hiển thị các biểu đồ và số liệu theo yêu cầu. Đặc biệt, sản phẩm tích hợp chatbot giúp người dùng dễ dàng tra cứu dữ liệu và thông tin. Tính năng hỏi đáp thông minh cho phép người dùng đặt câu hỏi trực tiếp bằng ngôn ngữ tự nhiên và nhận lại câu trả lời trong thời gian dưới 30 giây.

Tính khả thi của sản phẩm được đảm bảo nhờ khả năng tự động hóa cao, giúp giảm đáng kể thời gian vận hành và chi phí bảo trì. Với thiết kế linh hoạt và khả năng mở rộng, sản phẩm có tiềm năng duy trì và phát triển lâu dài, đáp ứng nhu cầu ngày càng phức tạp của người dùng. Người dùng được hỗ trợ không chỉ là các nhà quản lý, chuyên gia làm việc trong lĩnh vực kinh tế mà còn cả trong lĩnh vực quản lý ngành năng lượng. Sản phẩm có thể hỗ trợ thu thập, phân tích dữ liệu vĩ mô liên quan đến thị trường năng lượng như: cung - cầu, xu hướng giá cả; tối ưu hóa quá trình quản lý dữ liệu và phân tích thị trường; xây dựng báo cáo phân tích các yếu tố ảnh hưởng đến ngành năng lượng cho các cơ quan quản lý nhà nước



Ngày nhận bài: 20/11/2024

Ngày đánh giá và sửa chữa: 20/11 - 18/12/2024

Ngày duyệt đăng: 18/12/2024.

và tổ chức nghiên cứu; các công ty tư vấn và tổ chức phân tích năng lượng...

Bài viết giới thiệu sản phẩm số phục vụ phân tích dữ liệu kinh tế vĩ mô tích hợp trí tuệ nhân tạo giúp tự động tìm kiếm và thu thập dữ liệu từ nhiều nguồn, bao gồm báo cáo kinh tế, dữ liệu từ các tổ chức quốc tế, dữ liệu tài chính từ các sàn giao dịch và các nguồn dữ liệu công khai khác.

2. Đánh giá các sản phẩm phân tích kinh tế vĩ mô

Các sản phẩm phân tích dữ liệu kinh tế trên thị trường cung cấp nguồn thông tin phong phú với khối lượng lớn, đáng tin cậy và được cập nhật định kỳ (theo ngày, tháng, quý hoặc năm). Tuy nhiên, khối lượng và tính cập nhật của dữ liệu trong các sản phẩm này phụ thuộc đáng kể vào mục tiêu sử dụng và nguồn lực của đơn vị phát triển. Các sản phẩm thường có ưu điểm trong việc cung cấp thông tin tổng hợp nhưng đôi khi gặp hạn chế về độ trễ trong cập nhật dữ liệu hoặc mức độ chi tiết phù hợp cho các nhu cầu phân tích sâu.

Cơ sở dữ liệu của Tổng cục Thống kê (GSO) là nguồn thông tin quan trọng và chính thức, cung cấp dữ liệu về các lĩnh vực kinh tế - xã hội của Việt Nam. Ưu điểm của nguồn dữ liệu này là tính chính xác cao và đáng tin cậy, với các chỉ số vĩ mô như GDP, lạm phát, tỷ lệ thất nghiệp, xuất nhập khẩu, dân số và các chỉ tiêu phát triển bền vững. Dữ liệu được phân nhóm, hỗ trợ người dùng tra cứu hoặc yêu cầu phân tích toàn diện và chuyên sâu, với khối lượng lớn và phong phú. Dữ liệu thường được công bố định kỳ theo tháng, quý hoặc năm, tùy thuộc vào loại chỉ tiêu, giúp duy trì tính ổn định và đáng tin cậy của thông tin. Tuy nhiên, nhược điểm của cơ sở dữ liệu này là chậm đáp ứng nhu cầu phân tích thời gian thực hoặc biến động ngắn hạn trên thị trường, do dữ liệu được công bố định kỳ thay vì liên tục.

Trong khi đó, Ngân hàng Thế giới (World Bank) cung cấp một khối lượng lớn dữ liệu về các chỉ số kinh tế vĩ mô và xã hội của hầu hết các quốc gia, với các lĩnh vực như phát triển bền vững, xóa đói giảm nghèo và giáo dục. Ưu điểm của nguồn dữ liệu này là phong phú và toàn diện, phù hợp với các phân tích toàn cầu và khu vực. Dữ liệu được cập nhật định kỳ hàng năm hoặc hàng quý, giúp cung cấp cái nhìn tổng quan về các vấn đề kinh tế và xã hội. Tuy nhiên, nhược điểm của cơ sở dữ liệu này cũng liên quan đến độ trễ trong việc cập nhật dữ liệu mới.

Ngược lại, Financial Times có khả năng cung cấp dữ liệu tài chính và kinh tế theo thời gian thực với khả năng

cập nhật liên tục các biến động trên thị trường toàn cầu. Ưu điểm của nguồn dữ liệu này là tính kịp thời và chi tiết, cho phép theo dõi các thay đổi thị trường tức thời, đặc biệt trong lĩnh vực chứng khoán, tiền tệ, hàng hóa... Khối lượng dữ liệu lớn và được phân nhóm, giúp người dùng dễ dàng truy cập và phân tích. Tuy nhiên, nhược điểm của nguồn dữ liệu này là phạm vi chủ yếu tập trung vào các yếu tố tài chính, dẫn đến hạn chế khi sử dụng trong phân tích dài hạn và toàn diện về tình hình kinh tế - xã hội.

Với xu hướng trực quan hóa dữ liệu như hiện nay, các sản phẩm phân tích kinh tế thường được biểu diễn thông qua các dashboard. Ưu điểm của các sản phẩm dashboard trên thị trường là sự đa dạng về tính năng, giao diện và mức độ thân thiện với người dùng, đáp ứng nhiều mục tiêu phát triển khác nhau của các tổ chức. Tính năng nổi bật bao gồm khả năng tùy chỉnh biểu đồ, truy vấn dữ liệu theo thời gian thực và tích hợp các công cụ phân tích chuyên sâu. Về giao diện, hầu hết các sản phẩm có thiết kế hiện đại, tập trung vào trải nghiệm người dùng với bố cục hợp lý và khả năng tương tác cao. Tuy nhiên, các sản phẩm trên thị trường vẫn tồn tại một số hạn chế liên quan đến độ thân thiện. Dashboard hướng đến người dùng phổ thông thường dễ tiếp cận nhưng giới hạn về chức năng tùy chỉnh. Trong khi đó, các công cụ chuyên sâu dành cho nhà phân tích hoặc nhà đầu tư đòi hỏi kỹ năng sử dụng nhất định, khiến các sản phẩm trở nên khó tiếp cận đối với người dùng không chuyên.

Ví dụ, infographic của GSO có ưu điểm là tập trung vào việc truyền tải thông tin dưới dạng biểu đồ, hình ảnh và số liệu trực quan, với các chỉ số được trình bày dễ hiểu, phù hợp với người dùng phổ thông. Giao diện của sản phẩm đơn giản, dễ tiếp cận, giúp người dùng nhanh chóng nắm bắt thông tin cơ bản. Tuy nhiên, hạn chế của sản phẩm là chưa tích hợp nhiều công cụ tương tác và phần lớn nội dung mang tính tĩnh, khó đáp ứng các nhu cầu tùy chỉnh hoặc khai thác dữ liệu linh hoạt của người dùng.

Ngược lại, dashboard của World Bank [5] hoặc Tổ chức Hợp tác và Phát triển Kinh tế (OECD) [6] có ưu điểm là cung cấp các công cụ trực quan hóa, cho phép tùy chỉnh biểu đồ và truy vấn dữ liệu theo quốc gia, khu vực và các chỉ số cụ thể. Giao diện được thiết kế hợp lý với hệ thống lựa chọn rõ ràng, dễ sử dụng, phù hợp cho cả người dùng không chuyên. Tuy nhiên, các dashboard này chưa tích hợp công cụ hỏi đáp ứng dụng AI, làm giảm khả năng hỗ trợ thông minh và hiệu quả đối với người dùng cần giải pháp có tính tương tác cao hơn. Ngoài ra, sự đa dạng trong các loại dashboard còn hạn chế khi chỉ có một số

chỉ số được biểu diễn, chưa đáp ứng được nhu cầu phân tích phong phú ở nhiều lĩnh vực khác nhau của nền kinh tế.

Tương tự, dashboard của Financial Times [7] có ưu điểm nổi bật ở khả năng tương tác, cho phép người dùng điều chỉnh biểu đồ theo thời gian hoặc danh mục đầu tư. Đây là công cụ hỗ trợ hữu ích cho việc phân tích thị trường và theo dõi liên tục biến động tài chính. Tuy nhiên, sản phẩm này cũng có hạn chế tương tự một số dashboard khác khi chưa tích hợp công cụ AI giúp tối ưu hóa trải nghiệm người dùng.

Những hạn chế trên đã mở ra nhu cầu cấp thiết cho việc tích hợp công nghệ mới như AI vào quy trình phân tích dữ liệu kinh tế, khi tốc độ và khả năng thích ứng là yếu tố then chốt.

3. Ứng dụng AI phát triển sản phẩm số phân tích kinh tế vĩ mô

Sản phẩm “Phân tích số ứng dụng AI/ML tổng hợp thông tin thị trường và dự báo chỉ tiêu kinh tế” được VPI triển khai trên nền tảng Power BI và Microsoft Fabric, đảm bảo cho việc bảo mật, lưu trữ và cộng tác trong quá trình xây dựng sản phẩm, khả năng tự động hóa và tái sử dụng đối với sản phẩm cũng như khả năng truy cập của người dùng cuối cùng.

Đối với tính năng tra cứu dữ liệu, sản phẩm gồm dashboard được xây dựng dựa trên các thông tin liên quan đến kinh tế thế giới và Việt Nam, phân bố tại các báo cáo khác nhau. Do đó, để tìm kiếm thông tin cụ thể, người dùng có thể truy cập vào trang báo cáo đã được đặt tên. Tại mỗi trang báo cáo đều bố trí các biểu đồ thể hiện thông tin khác nhau, người dùng có thể truy cập ngay. Hoặc nhờ vào khả năng liên kết dữ liệu và hiển thị nâng cao của Power BI, người dùng có thể trở con trỏ chuột vào phân vùng biểu đồ mong muốn và thông tin liên quan sẽ được hiển thị, gồm: kinh tế toàn cầu, kinh tế Việt Nam, thông tin thị trường và trợ lý ảo.

Đối với tính năng tùy chỉnh xem báo cáo, sản phẩm giúp người dùng lựa chọn kiểu xem (view) báo cáo theo khoảng thời gian nhất định, theo một hoặc nhiều nguồn thông tin. Ngoài ra, sản phẩm được thiết kế để người dùng có thể

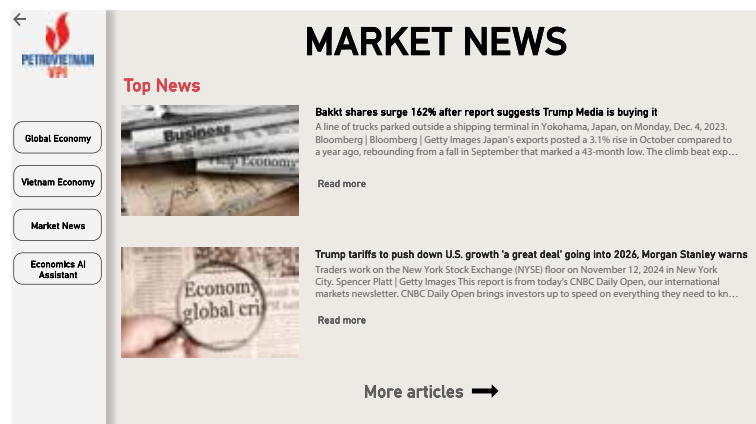
lựa chọn xem theo: thông tin tóm tắt, thông tin đầy đủ và các thông tin liên quan khác...

Đối với tính năng hỏi đáp với trợ lý ảo, sản phẩm cho phép người dùng đặt các câu hỏi liên quan đến thị trường kinh tế ngoài phạm vi dữ liệu được lưu trữ của sản phẩm. Tính năng này được xây dựng dựa trên trí tuệ nhân tạo AI, giúp nhận biết các yêu cầu tra cứu của người dùng.

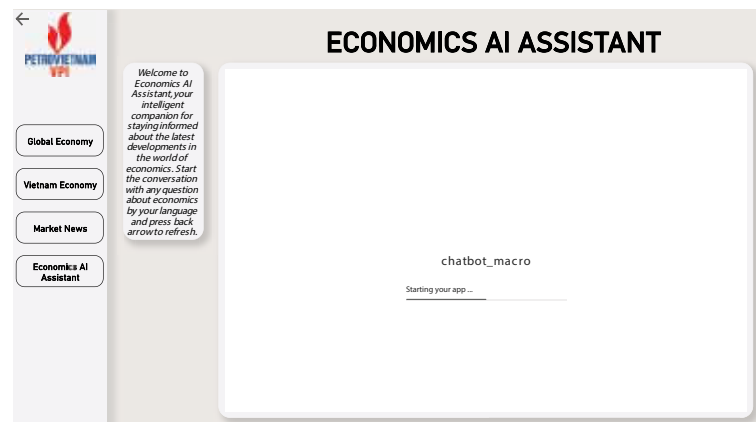
Về giao diện, cấu trúc của giao diện sản phẩm được tạo bởi các trang báo cáo; mỗi trang báo cáo có các biểu đồ tương ứng với tính năng hoặc nhóm thông tin có liên quan với nhau về nền kinh tế.



Hình 1. Giao diện điều hướng báo cáo.



Hình 2. Tính năng điều chỉnh view và xem báo cáo.

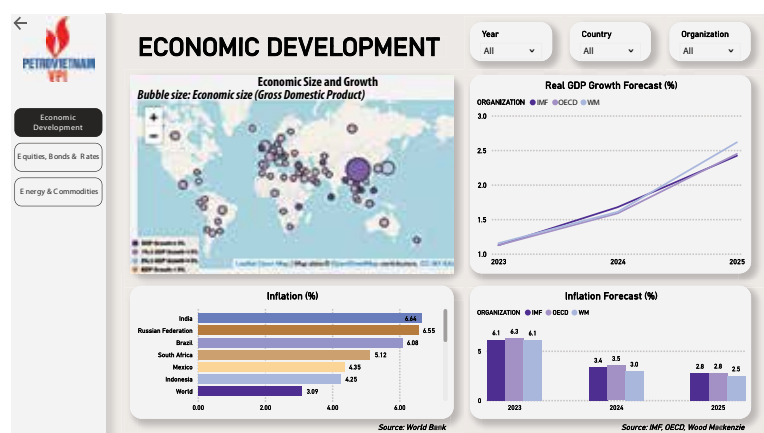


Hình 3. Tính năng hỏi đáp với trợ lý ảo.

Cấu trúc giao diện gồm 6 trang báo cáo: Economic Development; Equities, Bonds & Rates; Energy & Commodities; Vietnam Economy; Market News và Economics AI Assistant.

Để xây dựng sản phẩm, dữ liệu cấu trúc và phi cấu trúc từ kho dữ liệu VPI và các tổ chức (World Bank, Wood Mackenzie, Reuters,...) được thu thập, tổng hợp và xử lý với Python Notebook trên nền tảng Microsoft Fabric phục vụ truy vấn dữ liệu tự động hóa tối đa. Một số đặc điểm dữ liệu được mô tả trong Bảng 1.

Với mục tiêu nâng cao hiệu quả tốc độ xử lý, chi phí duy trì và khả năng tái sử dụng của sản phẩm, AI đã được ứng dụng trong một số quy trình kỹ thuật.



Hình 4. Giao diện biểu diễn các chỉ tiêu tăng trưởng kinh tế.

- Tổng hợp và xử lý thông tin

Các công cụ AI như Exa [8] và Cohere [9] được triển khai để tự động hóa các quy trình này, giảm thiểu sự can thiệp của con người và tối ưu hóa độ chính xác của dữ liệu thu thập được. Exa đảm nhận việc thu thập và xử lý dữ liệu thô, trong khi Cohere chịu trách nhiệm tổng hợp và tóm tắt thông tin.

Để đảm bảo chất lượng dữ liệu đầu vào, nhóm tác giả dựa vào kiến thức chuyên môn nhằm lựa chọn những nguồn thông tin uy tín: Reuters, CNN... và thiết lập các tham số tìm kiếm bao gồm: domain (nguồn), category (danh mục), query (truy vấn) và ngày cập nhật. Trong phạm vi sản phẩm này, các nguồn cung cấp tin như Reuters, CNN..., danh mục "news" (tin tức) và ngày cập nhật gần nhất được ưu tiên đưa vào tìm kiếm. Cách tiếp cận này đảm bảo dữ liệu đầu vào liên quan trực tiếp đến chủ đề phân tích, có tính kịp thời và độ chính xác cao.

Thông tin được thu thập đảm bảo không chứa các yếu tố nhiễu hoặc không liên quan đến chủ đề của sản phẩm. Để giảm thiểu kết

Bảng 1. Mô tả đặc điểm các dữ liệu trong cơ sở dữ liệu

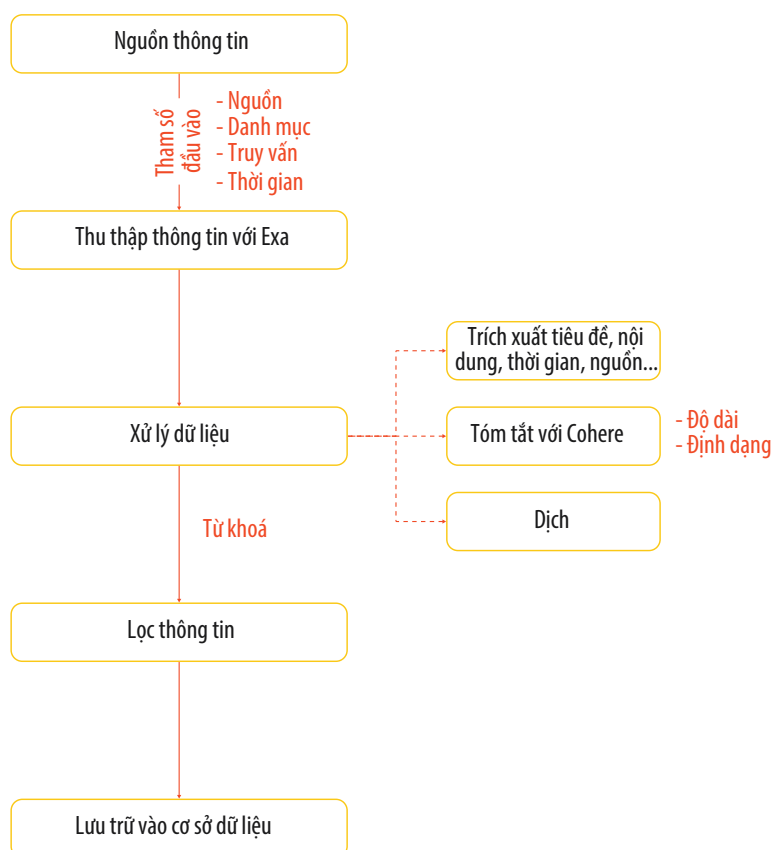
Chỉ số/chỉ tiêu	Nguồn	Giai đoạn	Tần suất
Tổng sản phẩm quốc nội (90 quốc gia)	World Bank	2010 - Hiện tại	Năm
Lạm phát (20 quốc gia)	World Bank	2010 - Hiện tại	Năm
Dự báo tổng sản phẩm quốc nội toàn cầu	IMF, OECD, Wood Mackenzie	2023 - 2025	Năm
Dự báo lạm phát toàn cầu	IMF, OECD, Wood Mackenzie	2023 - 2025	Năm
Lãi suất các Ngân hàng Trung ương (Mỹ, Anh, châu Âu)	Cục Dự trữ Liên bang, Ngân hàng Trung ương Anh - châu Âu	2018 - Hiện tại	Tháng
Lợi suất trái phiếu Chính phủ Mỹ 2 năm	Cục Dự trữ Liên bang	2018 - Hiện tại	Ngày
Lợi suất trái phiếu Chính phủ Mỹ 10 năm	Cục Dự trữ Liên bang	2018 - Hiện tại	Ngày
Chỉ số MSCI toàn cầu	Investing	2018 - Hiện tại	Ngày
Chỉ số MSCI châu Á (trừ Nhật Bản)	Yahoo!Finance	2018 - Hiện tại	Ngày
Giá dầu thô Dated Brent	EIA	2018 - Hiện tại	Ngày
Giá khí tự nhiên	EIA	2018 - Hiện tại	Ngày
Giá vàng	Investing	2018 - Hiện tại	Ngày
Giá hợp đồng tương lai dầu thô Dated Brent	ICE	2023 - Hiện tại	Giờ
Giá hợp đồng tương lai khí tự nhiên	CME	2023 - Hiện tại	Giờ
Giá hợp đồng tương lai vàng	CME	2023 - Hiện tại	Giờ
Chỉ số quản lý thu mua Việt Nam	S&P Global	2018 - Hiện tại	Tháng
Chỉ số giá tiêu dùng Việt Nam	Tổng cục Thống kê	2018 - Hiện tại	Tháng
Tỷ giá hối đoái Việt Nam đồng với USD, bảng Anh và EUR	Investing	2018 - Hiện tại	Ngày
Thông tin thị trường	Reuters, CNN...	2023 - Hiện tại	Ngày

quả nhiều, nhóm tác giả đã thiết kế query cụ thể, tập trung vào nội dung liên quan đến chủ đề muốn tìm kiếm. Các query này được xây dựng với một số từ khóa và điều kiện nhằm định hướng hệ thống chỉ thu thập những thông tin cần thiết. Ví dụ query được sử dụng với Exa: Các tin tức kinh tế mới nhất bao gồm, nhưng không giới hạn trong các nội dung về tài chính, tiền tệ, GDP, lạm phát, lãi suất, triển vọng kinh tế.

Query này không chỉ giới hạn phạm vi thu thập vào các lĩnh vực quan trọng trong nền kinh tế mà còn lọc bỏ các nội dung ít liên quan hoặc không phù hợp ngay từ bước đầu tiên. Việc sử dụng query cụ thể này đảm bảo dữ liệu đầu vào đã qua bước sàng lọc sơ bộ, giảm thiểu khối lượng thông tin nhiễu phải xử lý trong các giai đoạn tiếp theo.

Thông tin thị trường kinh tế được xử lý và tóm tắt bằng công cụ xử lý ngôn ngữ tự nhiên Cohere. Để hướng dẫn Cohere xử lý thông tin chính xác và hiệu quả, nhóm tác giả đã cung cấp tham số về độ dài, định dạng. Đối với sản phẩm này, Cohere được yêu cầu tóm tắt văn bản với độ dài trung bình (medium) và định dạng đoạn (paragraph). Thông tin sẽ được dịch qua tiếng Việt với các thư viện cung cấp khả năng dịch văn bản trên ngôn ngữ lập trình Python.

Sau khi được xử lý và tóm tắt bằng Cohere, thông tin tiếp tục được kiểm tra để đánh giá nhằm nâng cao chất lượng. Để đảm bảo thông tin đưa vào báo cáo số thuộc chủ đề kinh tế, nhóm tác giả sử



Hình 5. Quy trình tổng hợp và xử lý thông tin.

dụng ngôn ngữ lập trình Python nhằm lọc tin theo các từ khóa cụ thể. Sau khi xác định các từ khóa liên quan đến chủ đề kinh tế, chẳng hạn như "GDP", "lạm phát", "thị trường chứng khoán", "chính sách tiền tệ", "chỉ số giá tiêu dùng (CPI)". . . , các từ khóa này được sử dụng để quét và phân loại thông tin dựa vào nội dung đã tóm tắt từ Cohere. Bước này đảm bảo báo cáo tập trung vào thông tin liên quan trực tiếp đến chủ đề kinh tế, giúp nâng cao chất lượng thông tin và độ chính xác với chủ đề.

Quy trình tổng hợp và xử lý thông tin ứng dụng công cụ AI giúp tiết kiệm thời gian và chi phí do giảm thiểu được nhân lực làm việc thủ công, đồng thời đảm bảo chất lượng thông tin do có sự kiểm soát của nhóm chuyên môn. Đồng thời, quy trình này có thể được sử dụng trong một số sản phẩm có yêu cầu tương tự, giúp tiết kiệm thời gian xây dựng mới.

- Chatbot hỏi đáp thông tin kinh tế

Chatbot hỗ trợ hỏi đáp thông tin thị trường kinh tế được xây dựng trên nền tảng Copilot Studio [10], cung cấp thông tin cập nhật về nền kinh tế toàn cầu và Việt Nam. Với việc ứng dụng các công nghệ tiên tiến như trí tuệ nhân tạo và xử lý ngôn ngữ tự nhiên (NLP), chatbot giúp người dùng dễ dàng tiếp cận và tìm hiểu về các chỉ số kinh tế, tin tức thị trường và các sự kiện tài chính quan trọng.

Để đảm bảo chatbot có thể tìm kiếm thông tin chính xác và phù hợp với câu hỏi của người dùng, kỹ thuật prompting đã được phát triển. Đây là tập hợp các đoạn văn bản hoặc chỉ thị cụ thể, thiết kế dựa trên kiến thức chuyên môn của nhóm tác giả trong lĩnh vực kinh tế và cấu trúc dữ liệu chuyên ngành. Các prompt này được xây dựng với mục tiêu định hướng chatbot trong việc tìm kiếm thông tin từ các nguồn đáng tin cậy và uy tín, đồng thời đảm bảo rằng thông tin được cung cấp sát với yêu cầu của người dùng. Ngoài ra, chatbot còn được thiết kế để tìm kiếm và phản hồi theo ngôn ngữ của người dùng, giúp nâng cao trải nghiệm người dùng và đảm bảo rằng thông tin nhận được là dễ hiểu và chính xác. Quá trình thử nghiệm và cải tiến các prompt với sự tham gia của nhóm chuyên môn và

người dùng cuối giúp nâng cao chất lượng của chatbot. Các prompt được điều chỉnh và cập nhật dựa trên phản hồi và kết quả thực tế, đảm bảo rằng chatbot có thể đáp ứng các yêu cầu đa dạng, hiệu quả hơn trong việc cung cấp thông tin kinh tế quan trọng.

Ưu điểm lớn nhất của việc tích hợp AI và NLP là khả năng tạo ra giao diện thân thiện và dễ sử dụng. Cụ thể, chatbot hỏi đáp sử dụng AI và NLP cho phép người dùng tương tác với hệ thống bằng ngôn ngữ tự nhiên mà không cần phải học các thao tác phức tạp. Bên cạnh đó, việc tích hợp AI và NLP là khả năng tái sử dụng và mở rộng hệ thống. Các chatbot được xây dựng trên nền tảng AI có thể dễ dàng cập nhật và mở rộng để đáp ứng nhu cầu thay đổi của người dùng mà không cần phải thiết kế lại toàn bộ hệ thống, đồng thời có thể được kết nối và sử dụng trên một sản phẩm trong lĩnh vực tương tự.

4. Đánh giá sản phẩm phân tích số ứng dụng AI/ML tổng hợp thông tin thị trường và dự báo chỉ tiêu kinh tế

Sản phẩm “Phân tích số ứng dụng AI/ML tổng hợp thông tin thị trường và dự báo chỉ tiêu kinh tế” có ưu điểm hơn so với một số sản phẩm khác trên thị trường.

- Độ chính xác của các mô hình AI được sử dụng giúp nâng cao chất lượng sản phẩm. Exa và Cohere là các mô hình được huấn luyện trên bộ dữ liệu lớn và đa dạng, được thiết kế với khả năng tự cải thiện thông qua cơ chế học liên tục, thường xuyên được tinh chỉnh và cập nhật dựa trên phản hồi của người dùng và xu hướng phát triển công nghệ. Việc ứng dụng các mô hình này vào sản phẩm giúp đảm bảo độ chính xác của việc thu thập, tổng hợp dữ liệu và thông tin liên quan đến chủ đề tới hơn 90%. Tính chính xác của các mô hình AI không chỉ được đảm bảo nhờ khả năng vượt trội của công nghệ mà còn được củng cố bởi sự can thiệp của nhóm chuyên môn dựa trên kết quả từ mô hình. Nhờ sự kết hợp này, sản phẩm có thể đạt độ chính xác cao hơn so với các giải pháp chỉ sử dụng AI hoặc chỉ dựa vào kiến thức chuyên môn của con người.

- Khả năng xử lý dữ liệu của sản phẩm có ưu điểm là tốc độ xử lý nhanh và quản lý khối lượng dữ liệu lớn. Ngôn ngữ lập trình cũng như các công cụ AI được sử dụng hỗ trợ tự động cập nhật thường xuyên và xử lý hiệu quả dữ liệu mới trong vài chục giây đến vài phút. Khả năng này giúp tiết kiệm thời gian thực hiện và giảm sự can thiệp thủ công so với một số giải pháp trên thị trường hiện nay.

- Chi phí triển khai và vận hành được tối ưu nhờ ứng dụng các công cụ và quy trình tích hợp AI. So với một số giải pháp yêu cầu đội ngũ nhân công lớn để thu thập, xử

lý và tổng hợp dữ liệu, sản phẩm này giúp tiết kiệm chi phí nhân công do tự động hóa tương đối các bước nêu trên.

- Các mô hình như Exa và Cohere hoạt động dựa trên cơ sở hạ tầng điện toán đám mây, giảm nhu cầu đầu tư ban đầu vào phần cứng. Với khả năng tự động hóa và tích hợp linh hoạt, sản phẩm giúp giảm thiểu chi phí đào tạo, thiết kế lại hoặc bổ sung nguồn lực khi mở rộng sang lĩnh vực khác, đồng thời đảm bảo tính nhất quán và khả năng mở rộng linh hoạt, dễ dàng tích hợp thêm nguồn dữ liệu hoặc mở rộng phạm vi phân tích.

Tuy nhiên, sản phẩm vẫn còn hạn chế khi phụ thuộc vào chất lượng dữ liệu đầu vào. Nếu dữ liệu đầu vào không đầy đủ, không đồng nhất, hoặc có sai sót, kết quả phân tích có thể không đáng tin cậy. Hơn nữa, AI đôi khi tạo ra những cảnh báo hoặc dự đoán khó hiểu hoặc khó xác thực, dẫn đến rủi ro trong việc ra quyết định. Do đó, việc xây dựng các cơ chế kiểm tra và đánh giá chất lượng dữ liệu cũng như kết quả phân tích là điều cần thiết để đảm bảo độ tin cậy của sản phẩm. Mặc dù AI có khả năng xử lý và phân tích khối lượng lớn dữ liệu, nhưng con người vẫn đóng vai trò quan trọng trong việc kiểm tra, đánh giá, diễn giải và bổ sung các khía cạnh mà công nghệ không thể xử lý hoàn chỉnh, ví dụ bối cảnh thực tế, hiểu biết trong lĩnh vực hoặc các yếu tố văn hóa - xã hội. Sự kiểm soát của con người đảm bảo AI không chỉ hoạt động chính xác về mặt kỹ thuật mà còn phù hợp với các mục tiêu thực tế và yêu cầu của công việc.

Trong tương lai, sản phẩm có thể được phát triển, nâng cao chất lượng thông qua áp dụng các công nghệ nâng cao, sử dụng mô hình học máy hiện đại và phức tạp hơn giúp dự báo một số chỉ tiêu kinh tế. Ngoài ra, sản phẩm cần tiếp tục cải thiện các cơ chế kiểm soát chất lượng dữ liệu, xây dựng hệ thống phản hồi tự động từ người dùng để điều chỉnh mô hình và tăng cường tính tương tác thông qua trợ lý ảo hoặc chatbot hỗ trợ hỏi đáp.

5. Kết luận

Sản phẩm “Phân tích số ứng dụng AI/ML trong tổng hợp thông tin thị trường và dự báo chỉ tiêu kinh tế” giúp nâng cao tốc độ và quy mô xử lý dữ liệu thông qua việc ứng dụng AI. Hệ thống có khả năng tự động đọc, trích xuất và tóm tắt tin tức, báo cáo và dữ liệu chính thống gần thời gian thực. Đặc biệt, công nghệ AI cho phép tổng hợp và diễn giải thông tin theo ngữ cảnh, tạo ra các bản tóm tắt có trích dẫn nguồn rõ ràng. Nhờ khả năng linh hoạt và tích hợp cao, sản phẩm có thể dễ dàng kết nối trực tiếp với các hệ thống BI, công cụ trực quan hóa

dữ liệu và tùy biến các báo cáo phân tích chuyên sâu phục vụ yêu cầu quản lý, nghiên cứu và ra quyết định. Tuy nhiên, việc ứng dụng AI cần được quản trị chặt chẽ để bảo đảm độ tin cậy và tính minh bạch của hệ thống; có cơ chế đánh giá chất lượng và kiểm thử định kỳ; duy trì, hiệu chỉnh và cập nhật mô hình thường xuyên để đảm bảo tính chính xác, ổn định và kịp thời của hệ thống phân tích.

Định hướng phát triển sản phẩm “Phân tích số ứng dụng AI/ML trong tổng hợp thông tin thị trường và dự báo chỉ tiêu kinh tế” tập trung vào tăng cường năng lực cốt lõi của hệ thống, đặc biệt là khả năng tương tác với người dùng và trích xuất tri thức chuyên sâu. Cụ thể, phát triển các mô hình AI có khả năng phân tích ngữ cảnh sâu sắc trong câu hỏi của người dùng, không chỉ dừng lại ở việc phân hồi đơn lẻ mà là quy trình phân tích toàn diện để đưa ra câu trả lời chính xác, có căn cứ và trích dẫn nguồn rõ ràng. Tiếp theo là tăng khả năng trích xuất insight, thông qua việc liên kết AI với kho tri thức vĩ mô có cấu trúc, đảm bảo từng kết luận phân tích, phát hiện bất thường hay giải thích giả định đều được minh chứng bằng nguồn dữ liệu xác thực. Mục tiêu của nhóm tác giả là xây dựng trợ lý phân tích thông minh, giúp tăng độ minh bạch và hiệu suất ra quyết định..

Tài liệu tham khảo

[1] Paulo Silva, Carolina Goncalves, Nuno Antunes, Marilia Curado, and Bogdan Walek, “Privacy risk assessment and privacy-preserving data monitoring”, *Expert Systems with Application*, Volume 200, 2022. DOI: 10.1016/j.eswa.2022.116867.

[2] Tohid Atashbar and Rui Aruhan Shi, “AI and macroeconomic modelling: Deep reinforcement learning

in an RBC model”, 24/2/2023. [Online]. Available: <https://www.imf.org/en/Publications/WP/Issues/2023/02/24/AI-and-Macroeconomic-Modeling-Deep-Reinforcement-Learning-in-an-RBC-model-530084>.

[3] Rodolfo C. Cavalcante, Rodrigo C. Brasileiro, Victor L.F. Souza, Jarley P. Nobrega, and Adriano L.I. Oliveira, “Computational intelligence and financial markets: A survey and future directions”, *Expert Systems with Applications*, Volume 55, pp. 194 - 211, 2016. DOI: 10.1016/j.eswa.2016.02.006.

[4] B. Shraavan Kumar and Vadlamani Ravi, “A survey of the applications of text mining in financial domain”, *Knowledge-Based Systems*, Volume 114, pp. 128 - 147, 2016. DOI: 10.1016/j.knosys.2016.10.003.

[5] World Bank, “PPI visualization dashboard”. [Online]. Available: <https://ppi.worldbank.org/en/visualization#sector=&status=&ppi=&investment=®ion=&ida=&income=&ppp=&mbd=&year=&excel=false&map=&header=true>.

[6] OECD, “Households' economic well-being: The OECD dashboard”. [Online]. Available: <https://www.oecd.org/en/data/dashboards/households-economic-well-being-the-oecd-dashboard.html>.

[7] Financial Times, “Markets data”. [Online]. Available: <https://markets.ft.com/data>.

[8] Exa. [Online]. Available: <https://exa.ai/>.

[9] Cohere. [Online]. Available: <https://cohere.com/>.

[10] Copilot Studio. [Online]. Available: <https://copilotstudio.microsoft.com/>.

AI APPLICATIONS IN MARKET INFORMATION COLLECTION AND MACROECONOMIC INDICATOR ANALYSIS

Dao Huong Giang, Do Hong Hanh

Vietnam Petroleum Institute (VPI)

Email: giang.dh@vpi.pvn.vn

Summary

Data analysis in economics is transitioning from traditional methods to AI-driven approaches, enabling automation in data collection and processing, as well as improving the ability to forecast market fluctuations, providing a foundation for more accurate business decision-making. This article presents the application of AI at the Vietnam Petroleum Institute (VPI) to develop a digital analytics solution for collecting market information, aimed at improving the efficiency of data integration and automated processing, forecasting economic indicators, and providing a virtual assistant to support interactive market information queries.

Key words: Macroeconomics, Artificial intelligence, market analysis, data processing, virtual assistant.