

# CHATGPT, GENERATIVE AI AND INVESTMENT ADVISORY

Let's Get Started →



# BỐI CẢNH VÀ CÂU HỎI ĐẶT RA

- **Bối cảnh:**
  - **AI đang thay đổi cách phân tích tài chính, nhưng phương pháp cũ còn hạn chế.**
  - **Generative AI (như ChatGPT) mang lại tiềm năng mới nhờ khả năng suy luận và linh hoạt**



Câu hỏi đặt ra

ChatGPT có thể dự đoán danh mục đầu tư hiệu quả hơn không?

# GIẢI THÍCH THUẬT NGỮ

## AI & Models

- Generative AI :** là loại AI có thể tạo nội dung mới dựa trên dữ liệu đã học
- Fine-tuning:** quá trình đào tạo mô hình AI nhằm cải thiện độ chính xác trong dự đoán.
- LLM :** mô hình AI có hàng tỷ tham số được huấn luyện trên lượng lớn dữ liệu văn bản
- ChatGPT:** là một LLM được tạo ra bởi OpenAI có khả năng suy luận có tính ứng dụng trong nhiều lĩnh vực

## Data & Finance

- Textual analysis:** kỹ thuật trích xuất thông tin từ văn bản
- Alpha (Trong tài chính):** lợi nhuận vượt trội so với thị trường. Alpha dương nghĩa là danh mục đầu tư tạo lợi nhuận cao hơn kỳ vọng
- Temperature trong AI:** Tham số kiểm soát mức độ sáng tạo của LLM



# ABSTRACT



**Tối ưu hoá mô hình ChatGPT bằng dữ liệu huấn  
luyện đặc biệt và điều chỉnh tham số**



**Nghiên cứu dựa vào 2 loại dữ liệu từ WSJ và thông  
báo chính sách từ chính phủ TQ**



**Hiệu quả hơn các phương pháp thông thường vì  
có thể tạo ra alpha dương**

**Kết quả: Danh mục đầu tư từ ChatGPT có alpha 3  
nhân tố hàng tháng lên đến 3%. ChatGPT tạo ra  
danh mục đầu tư có hiệu suất cao nhất khi xử lý  
tin tức liên quan đến các chính sách kinh tế hoặc  
tài chính của chính phủ.**

# PHƯƠNG PHÁP TRUYỀN THỐNG TRONG ĐẦU TƯ

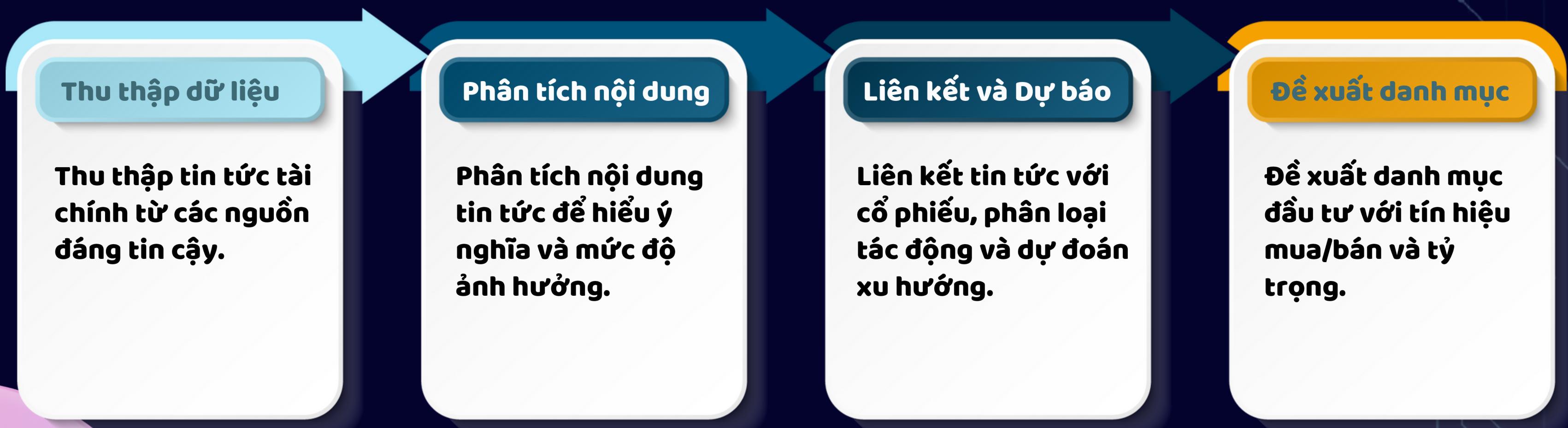
- **TF-IDF:** Chỉ xác định tầm quan trọng của một từ trong văn bản mà không hiểu ngữ cảnh tổng thể (chỉ xác định tần suất xuất hiện của các từ quan trọng nhưng không hiểu)
- **Sentiment Analysis:** Chỉ phân loại thông tin dựa trên mức độ tích cực hoặc tiêu cực mà không xem xét mối liên hệ giữa các từ (xác định từ với cảm xúc tích cực hay tiêu cực, nó sẽ đánh giá bài báo này có tác động tốt hay xấu tới thị trường)
- **NLP cơ bản:** Dựa vào các mô hình thống kê và không có khả năng hiểu nội dung như LLM



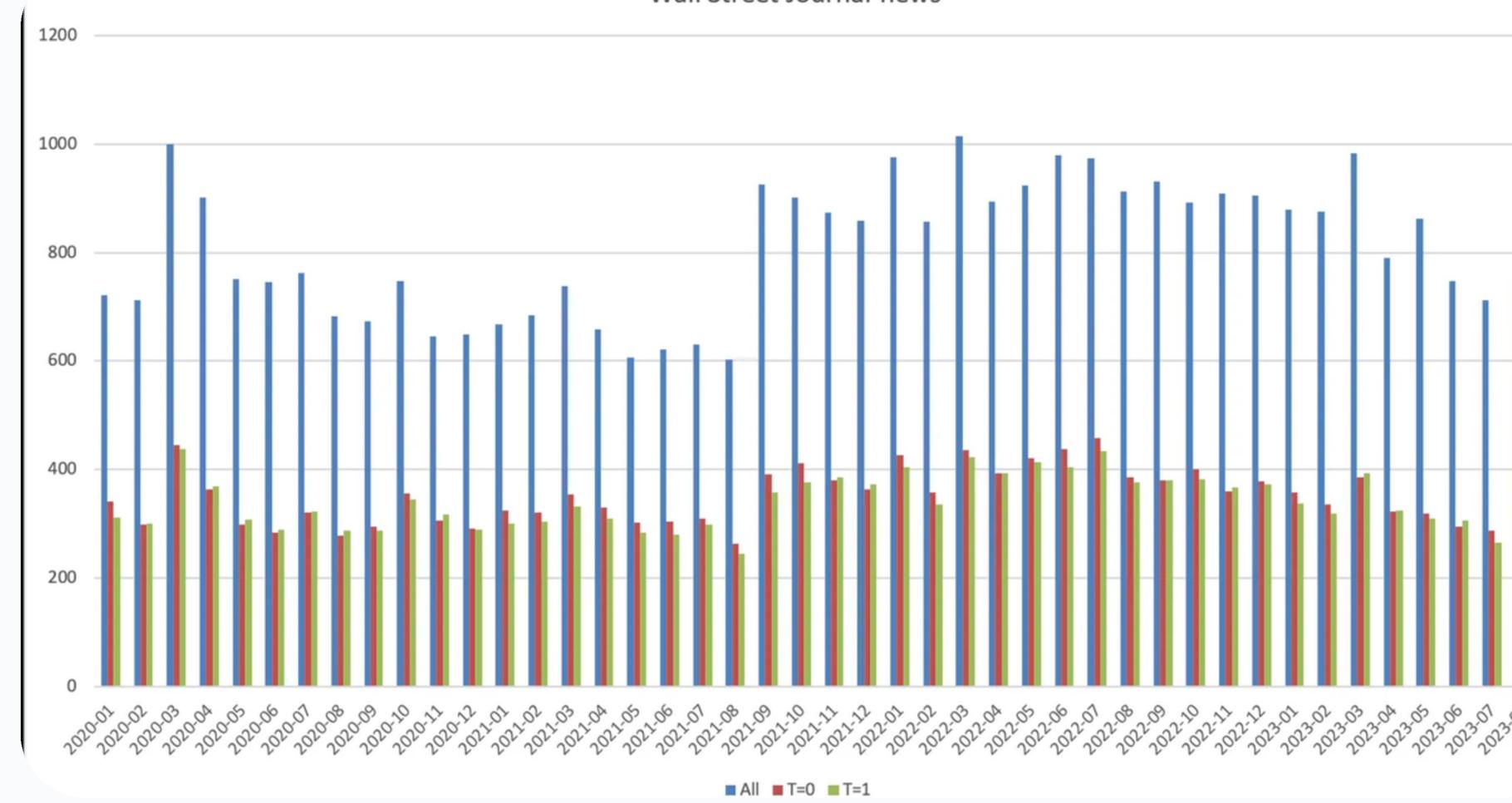
## Hạn chế

Không hiểu ngữ cảnh,  
chỉ dựa vào tần suất từ  
khóa

# CHATGPT TRONG PHÂN TÍCH ĐẦU TƯ

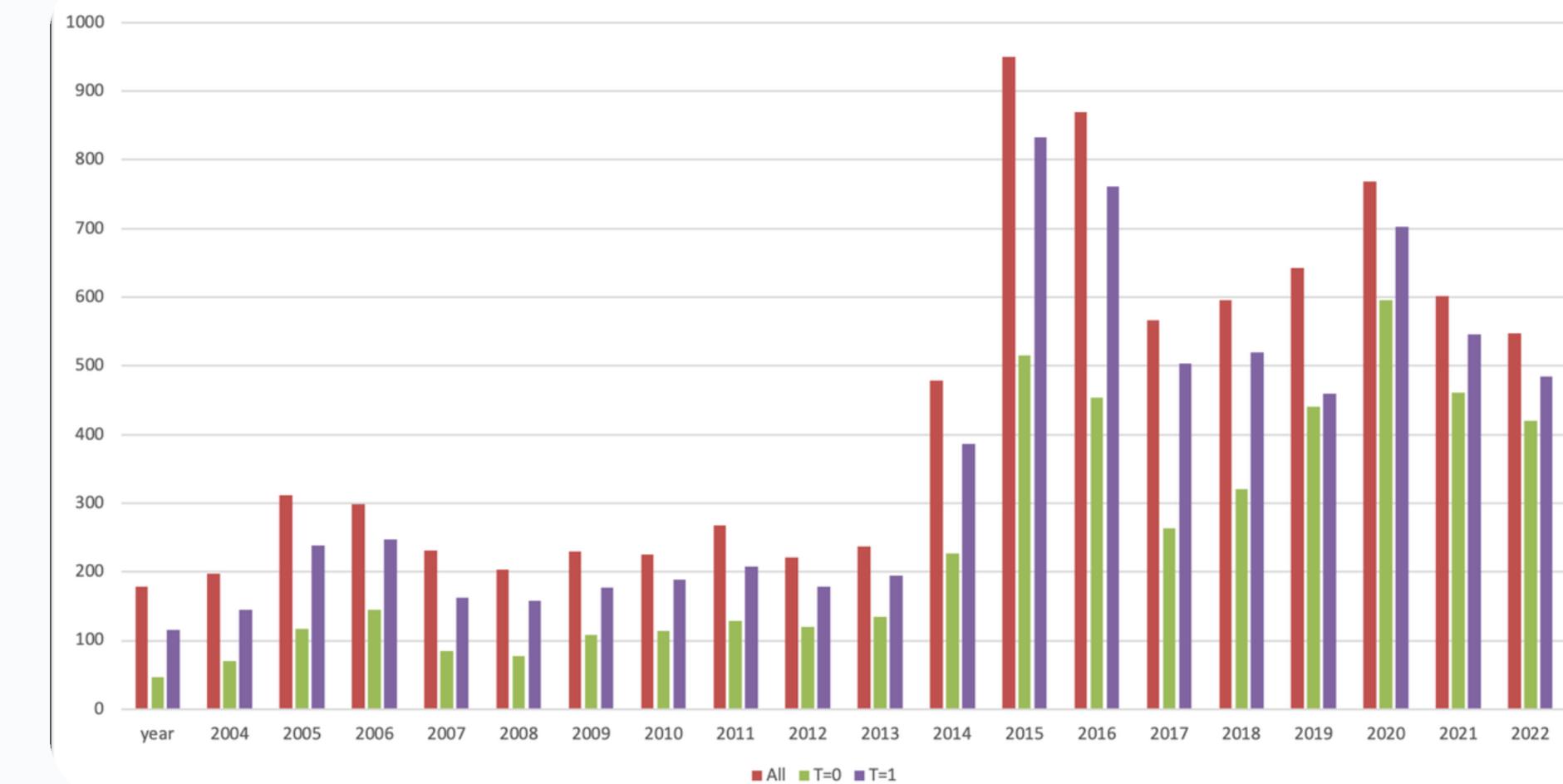


ChatGPT hiểu ngữ cảnh, linh hoạt không cần cấu trúc cố định, và có tiềm năng dự đoán chính xác hơn phương pháp truyền thống.



# DỮ LIỆU SỬ DỤNG TRONG NGHIÊN CỨU

China Policy News Release

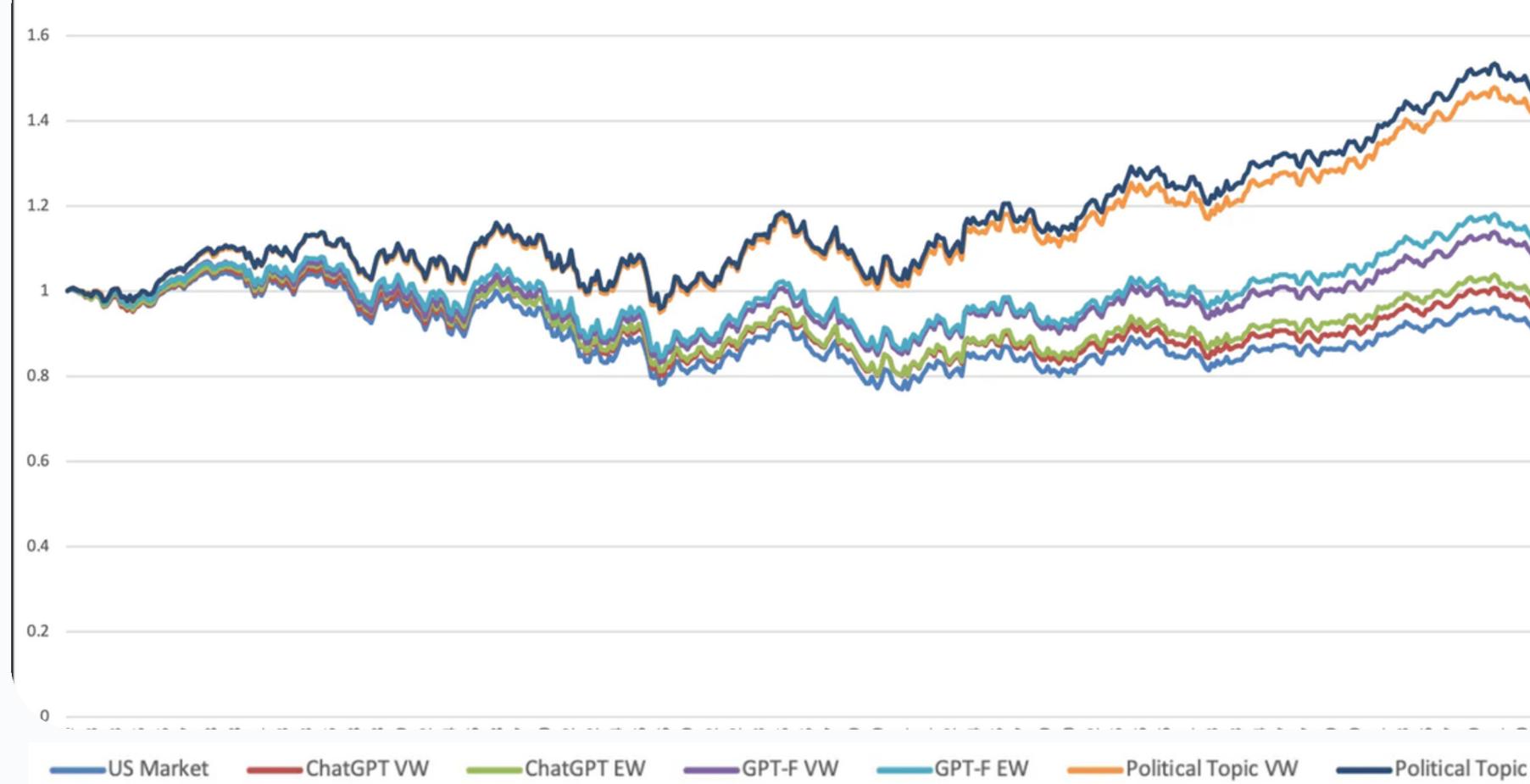


- Nguồn dữ liệu: WSJ (Mỹ) vs Chính sách (Trung Quốc)
- Thống kê về dữ liệu (số lượng bài báo ...)
- Biểu đồ cho thấy lượng bài báo hàng tháng của WSJ và số lượng báo cáo phát hành chính sách của TQ hàng năm.
  - T = 0 mô hình có xu hướng tạo ra kết quả ổn định, ít ngẫu nhiên (các bài báo liên quan đến khuyến nghị cổ phiếu).
  - T = 1 mô hình tạo ra kết quả đa dạng, sáng tạo hơn (các bài báo liên quan đến khuyến nghị cổ phiếu).

# PHƯƠNG PHÁP HUẤN LUYỆN MÔ HÌNH

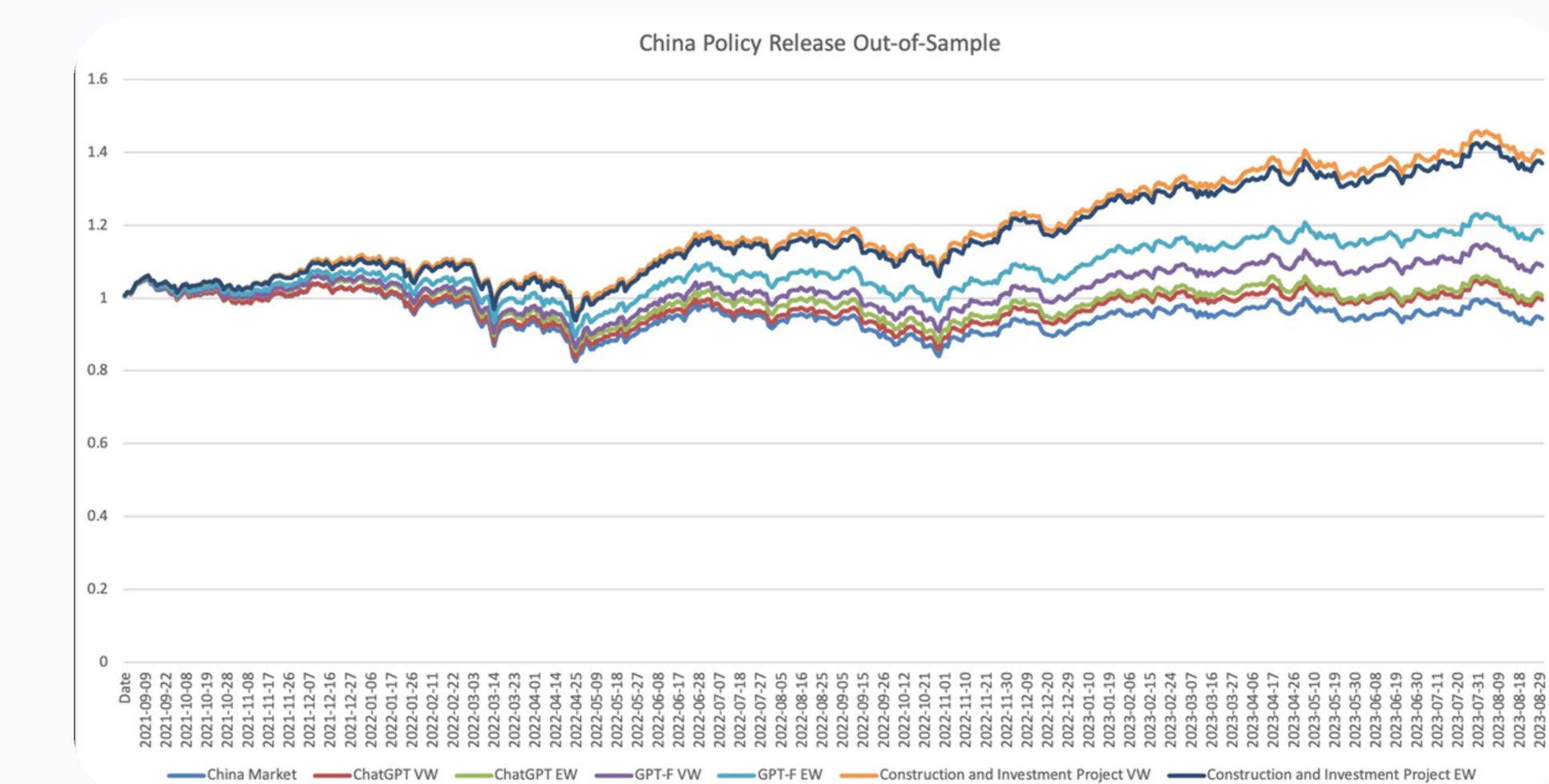
- **Truy vấn đàm thoại đơn giản:** hỏi đáp với ChatGPT về lĩnh vực cần nghiên cứu
- **Fine-tuning:** các nhà nghiên cứu đã tinh chỉnh trên một dữ liệu lớn về tin tức tài chính và thông báo chính sách
- Thuê các chuyên gia trong ngành và các sinh viên ngành tài chính để xem xét và đánh giá phản hồi của mô hình xem các khuyến nghị của ChatGPT có hợp lý và phù hợp không.
- **Thống kê mô tả:** tập trung vào việc phân tích đặc điểm của dữ liệu được sử dụng trong nghiên cứu.
- **Đánh giá khả năng tổng quát hoá của mô hình bằng In-sample test và Out-of-Sample tests.**
  - **In-sample test:** kiểm tra độ khớp với dữ liệu huấn luyện, nhưng có nguy cơ overfitting.
  - **Out-of-sample test:** đánh giá khả năng dự đoán thực tế, giúp kiểm chứng mô hình có thực sự hiệu quả không.
- **Kiểm tra tính mạnh mẽ của mô hình bằng cách sử dụng các tập dữ liệu khác nhau (In-sample và Out-of-sample), các phương pháp đánh giá (Các chỉ số tài chính, so sánh hiệu suất với các mô hình truyền thống) và các thông số mô hình khác nhau (thay đổi T).**

# HIỆU SUẤT ĐẦU TƯ CỦA CHATGPT



- **Kết luận:**

- **ChatGPT có thể nhận diện cơ hội đầu tư dựa trên tin tức**
- **Hiệu suất của danh mục đầu tư phụ thuộc vào loại thông tin được sử dụng**
- **Danh mục đầu tư của ChatGPT hoạt động tốt hơn GPT-F**
- **Thị trường biến động ảnh hưởng đến tất cả danh mục đầu tư**
- **ChatGPT có thể hỗ trợ quyết định đầu tư, nhưng không thay thế con người**



			CAPM	FF3	FF5	FF5 MOM+REV	
One-Month	ChatGPT 3.5	VW	T=0	3.592	3.403	3.621	3.383
			T=1	14.429**	14.104**	12.920*	13.853**
EW			T=0	4.618	4.503	4.415	4.516
			T=1	(0.614)	(0.593)	(0.583)	(0.605)
ChatGPT 4.0		VW	T=0	16.251**	15.104**	14.791**	13.560**
			T=1	(2.421)	(2.254)	(2.207)	(2.023)
EW			T=0	2.184	2.018	1.847	1.798
			T=1	(0.340)	(0.315)	(0.294)	(0.281)
		EW	T=0	1.618	1.384	1.231	1.029
			T=1	(0.243)	(0.201)	(0.183)	(0.142)
			T=0	1.729	1.492	1.320	1.283
			T=1	(0.257)	(0.231)	(0.231)	(0.197)
			T=0	1.537	1.321	1.298	1.252
			T=1	(0.229)	(0.194)	(0.186)	(0.181)

# TABLE 3 VÀ TABLE 4 (US - IN-SAMPLE, JAN 2020 - AUG 2021)

- **Nhận xét**
  - **Kỳ 1 tháng có alpha cao nhất (ChatGPT 3.5, T=1: 14.429 bps cho VW CAPM).**
  - **ChatGPT 3.5 vượt trội ChatGPT 4.0 ở hầu hết các kỳ.**
- **Table 4: Phân tích yếu tố (factor exposure) cho kỳ 1 tháng.**
- **Nhận xét:**
  - **Alpha cao nhất khi T=1: 13.853 bps (VW, FF5+MOM&REV), 15.104 bps (EW, FF3).**
  - **Các yếu tố như MKT, SMB có ý nghĩa thống kê cao, nhưng MOM và REV không rõ ràng.**

	FF3 Factors	FF5 Factors +MOM&REV	FF3 Factors	FF5 Factors +MOM&REV
	VW	VW	EW	EW
(b) Temperature=1				
Alpha	14.104** (2.084)	13.853** (2.043)	15.104** (2.254)	13.560** (2.023)
MKT	0.859*** (4.276)	0.759*** (3.175)	1.095*** (5.085)	1.163*** (5.165)
SMB	-0.749*** (-3.391)	-0.683*** (-3.184)	0.326*** (4.193)	0.409*** (3.194)
HML	0.184 (1.482)	0.189 (1.501)	-0.296 (-1.184)	-0.291 (-1.325)
RMW		0.193 (0.284)		0.819 (1.039)
CMA			0.391*** (3.384)	0.104 (0.812)
MOM			0.165 (0.257)	-0.487 (-0.957)
REV			-0.376 (-0.847)	0.578* (1.859)
Portfolio FE	Y	Y	Y	Y
R-squared	0.742	0.751	0.831	0.823

# TABLE 9 (CHINA - IN-SAMPLE, JAN 2004 - AUG 2021)

	FF3 Factors VW	FF5 Factors +MOM&REV VW	FF3 Factors EW	FF5 Factors +MOM&REV EW
(b) Temperature=1				
Alpha	6.221*** (2.750)	6.982*** (3.187)	8.513*** (4.048)	8.511*** (4.001)
MKT	0.827*** (11.295)	0.804*** (10.587)	0.904*** (13.920)	0.901*** (13.159)
SMB	-0.556*** (-39.281)	-0.471*** (-17.815)	-0.058*** (-3.096)	-0.045* (-1.866)
HML	0.119*** (5.623)	0.264*** (7.323)	0.132*** (4.345)	0.167*** (5.072)
RMW		0.438*** (10.166)		-0.008 (-0.190)
CMA		-0.147*** (-2.841)		-0.140*** (-2.975)
MOM		0.940 (0.841)		0.495 (0.317)
REV		-0.188 (-0.581)		-0.450 (-1.077)
R-squared	0.756	0.761	0.795	0.801
Portfolio FE	Y	Y	Y	Y
Observations	4250	4250	4250	4250

	FF3 Factors VW	FF5 Factors +MOM&REV VW	FF3 Factors EW	FF5 Factors +MOM&REV EW
(a) Temperature=0				
Alpha	2.501 (1.280)	2.731 (1.365)	3.092 (1.530)	2.980 (1.426)
MKT	0.829*** (8.391)	0.835*** (7.816)	0.843*** (8.281)	0.856*** (9.902)
SMB	-0.247*** (-13.014)	-0.275*** (-8.306)	0.239*** (13.519)	0.188*** (5.867)
HML	0.004 (0.149)	-0.141*** (-3.113)	-0.047* (-1.772)	-0.039*** (-0.889)
RMW		0.476*** (8.866)		-0.115** (-2.207)
CMA			0.074 (1.144)	-0.017 (-0.268)
MOM			-0.582 (-0.839)	-0.281 (-0.374)
REV			0.832 (1.194)	0.840 (0.953)
R-squared	0.621	0.641	0.686	0.691

- Nhận xét:

- Kỳ 1 tháng là tối ưu ở cả hai thị trường, nhưng alpha ở Mỹ cao hơn Trung Quốc (14-15 bps vs. 6-8 bps).
- ChatGPT hoạt động tốt hơn khi T=1, cho thấy tính sáng tạo cải thiện hiệu suất trong giai đoạn in-sample.

# SO SÁNH HIỆU SUẤT (US OUT-OF-SAMPLE TABLE 5, TABLE 6)

FF5+MOM and REV				
	T=0	T=1		
	VW	EW	VW	
All	0.937 (0.148)	0.748 (0.112)	1.329 (0.211)	1.425 (0.248)
Topic1	4.136 (0.617)	3.104 (0.461)	2.140 (0.331)	3.185 (0.493)
Topic2	-0.315 (-0.056)	-0.215 (-0.031)	-0.524 (-0.074)	-0.498 (-0.063)
Topic3	7.314 (1.125)	6.914 (1.284)	11.931** (1.988)	12.683** (2.112)
Topic4	1.349 (0.214)	1.215 (0.221)	1.127 (0.198)	1.049 (0.181)
Topic5	-0.315 (-0.045)	-0.525 (-0.082)	0.817 (0.142)	0.752 (0.109)
Topic6	3.194 (0.447)	2.896 (0.424)	0.471 (0.089)	0.829 (0.120)
Topic7	-2.348 (-0.384)	-1.849 (-0.245)	1.431 (0.249)	1.109 (0.140)
Topic8	0.381 (0.063)	0.521 (0.080)	1.152 (0.189)	2.427 (0.384)
Topic9	0.371 (0.053)	0.314 (0.041)	-0.584 (-0.093)	-0.616 (-0.095)
Topic10	-0.821 (-0.107)	-0.742 (-0.096)	-1.482 (-0.214)	-1.315 (-0.208)
Topic11	0.942 (0.142)	1.455 (0.168)	3.175 (0.441)	3.821 (0.571)
Topic12	1.429 (0.197)	2.394 (0.378)	-0.285 (-0.034)	-0.681 (-0.123)
Topic13	0.281 (0.057)	0.521 (0.089)	2.436 (0.384)	4.215 (0.659)

FF5+MOM and REV				
	T=0	T=1		
	VW	EW	VW	
All	1.349 (0.242)	1.294 (0.228)	3.480 (0.546)	4.195 (0.705)
Topic1	4.481 (0.715)	3.914 (0.652)	2.251 (0.425)	2.742 (0.471)
Topic2	-0.250 (-0.038)	-0.425 (-0.072)	2.194 (0.405)	2.350 (0.420)
Topic3	7.481 (1.125)	7.039 (1.028)	15.429** (2.494)	15.402** (2.356)
Topic4	0.732 (0.134)	0.385 (0.064)	12.584** (2.247)	12.942** (2.314)
Topic5	0.415 (0.074)	0.632 (0.104)	2.921 (0.632)	3.194 (0.805)
Topic6	2.948 (0.525)	3.215 (0.659)	1.492 (0.255)	1.341 (0.245)
Topic7	0.435 (0.074)	0.671 (0.114)	6.745 (1.148)	7.931 (1.341)
Topic8	-0.149 (-0.028)	-0.482 (-0.084)	0.482 (0.058)	1.956 (0.241)
Topic9	1.491 (0.240)	1.501 (0.229)	1.462 (0.285)	0.728 (0.147)
Topic10	-0.452 (-0.071)	-0.842 (-0.154)	-0.457 (-0.073)	-0.741 (-0.110)
Topic11	3.942 (0.682)	3.526 (0.653)	11.436** (2.106)	11.841** (2.194)
Topic12	1.905 (0.310)	2.149 (0.338)	-0.743 (-0.114)	-0.859 (-0.147)
Topic13	1.145 (0.240)	1.942 (0.294)	12.195** (2.219)	12.104** (2.056)

- **Topic 3 (Politics): Alpha cao nhất (11.931-12.683 bps khi T=1).**
  - Alpha tổng thể thấp (0.748-1.425 bps), không ý nghĩa thống kê.
- **Table 6 (Fine-tuning):**
  - Alpha cải thiện rõ rệt: Topic 3 (15.429 bps), Topic 4 (12.584 bps), Topic 13 (12.195 bps).
  - Alpha tổng thể tăng lên 3.480-4.195 bps.

# TABLE 10 VÀ TABLE 11 (CHINA - OUT-OF- SAMPLE, SEP 2021 - FEB 2024)

FF5+MOM and REV				
	T=0	T=1		
	VW	EW	VW	
All	0.285 (0.053)	0.194 (0.028)	2.149 (0.481)	2.245 (0.490)
Topic1	2.194 (0.347)	2.581 (0.419)	11.737* (1.928)	12.572** (2.101)
Topic2	-1.482 (-0.217)	-0.910 (-0.148)	0.319 (0.042)	0.505 (0.083)
Topic3	0.752 (0.121)	0.281 (0.049)	7.294 (1.142)	7.968 (1.394)
Topic4	0.124 (0.028)	0.475 (0.089)	4.412 (0.698)	4.712 (0.751)
Topic5	1.304 (0.272)	1.472 (0.307)	0.317 (0.064)	0.589 (0.091)
Topic6	-0.771 (-0.142)	-0.429 (-0.083)	-0.736 (-0.148)	-0.831 (-0.170)
Topic7	0.385 (0.074)	0.471 (0.081)	1.304 (0.242)	1.523 (0.284)

	FF5+MOM and REV			
	T=0	T=1	VW	EW
All	0.582 (0.103)	0.492 (0.089)	3.185 (0.591)	4.528 (0.841)
Topic1	2.753 (0.409)	3.410 (0.527)	13.482** (2.294)	13.096** (2.153)
Topic2	-1.234 (-0.281)	-0.849 (-0.190)	1.489 (0.274)	2.186 (0.391)
Topic3	0.890 (0.125)	1.149 (0.272)	12.280** (2.149)	12.136** (2.109)
Topic4	0.741 (0.143)	0.530 (0.079)	10.210* (1.838)	10.109* (1.751)
Topic5	0.753 (0.142)	0.938 (0.194)	0.821 (0.174)	0.481 (0.068)
Topic6	0.148 (0.029)	0.228 (0.042)	11.295* (1.934)	11.482** (1.990)
Topic7	0.174 (0.032)	0.310 (0.042)	-0.851 (-0.127)	-0.472 (-0.073)

- **Table 11 (Fine-tuning):**
  - **Alpha Topic 1 tăng lên 13.482 bps, Topic 3 (12.280 bps), Topic 4 (10.210 bps).**
  - **Alpha tổng thể tăng lên 3.185-4.528 bps.**
- **Table 10: Alpha tổng thể thấp (2.149-2.245 bps), nhưng Topic 1 (Government investment) nổi bật (11.737-12.572 bps).**

# ĐƯA RA NHẬN XÉT



1

**Fine-tuning cải thiện đáng kể alpha ở cả hai thị trường, đặc biệt với tin tức chính sách (Politics ở Mỹ, Government investment ở Trung Quốc).**

2

**Mỹ có nhiều chủ đề sinh alpha hơn (Topic 3, 4, 11, 13), trong khi Trung Quốc tập trung vào Topic 1, 3, 4, 6.**

# SO SÁNH GIỮA PHƯƠNG PHÁP TRUYỀN THỐNG VÀ CHATGPT

- **Table 7 (US): Phương pháp cosine similarity không tạo alpha đáng kể (0.310-1.294 bps), thậm chí âm ở Topic 13.**
- **Table 12 (China): Tương tự, alpha âm hoặc không ý nghĩa (-3.149 đến 1.742 bps).**
- **Nhận xét: ChatGPT vượt trội hoàn toàn so với textual analysis truyền thống ở cả hai thị trường, đặc biệt sau fine-tuning.**

**ChatGPT tạo danh mục có alpha dương, phương pháp cũ thì không**

	FF5+MOM and REV			
	Cosine Similarity Annual Reports	Cosine Similarity Earning Calls	VW	EW
All	0.375 (0.063)	0.632 (0.108)	0.421 (0.073)	0.310 (0.053)
Topic3	1.294 (0.205)	-1.526 (-0.302)	0.928 (0.153)	0.740 (0.105)
Topic4	-0.610 (-0.094)	0.219 (0.031)	0.801 (0.129)	0.694 (0.104)
Topic11	-0.219 (-0.032)	-0.314 (-0.049)	-0.493 (-0.077)	-0.512 (-0.086)
Topic13	-1.184 (-0.211)	-1.592 (-0.270)	-2.316 (-0.372)	-2.210 (-0.330)

	FF5+MOM and REV			
	T=0	T=1	VW	EW
All	-0.729 (-0.124)	-1.250 (-0.201)	-0.976 (-0.151)	-1.025 (-0.164)
Topic1	-1.190 (-0.204)	-1.115 (-0.212)	-3.018 (-0.502)	-3.149 (-0.517)
Topic3	-1.493 (-0.218)	-1.841 (-0.308)	-2.491 (-0.405)	-3.144 (-0.501)
Topic4	1.039 (0.151)	1.742 (0.301)	0.728 (0.124)	0.931 (0.165)
Topic6	-0.537 (-0.089)	-0.626 (-0.105)	-1.416 (-0.231)	-2.256 (-0.351)

# KHẢ NĂNG HIỂU NGỮ CẢNH CỦA CHATGPT

Panel A: Financial Analyst			
ChatGPT 3.5			
	Domestic Equity (T=0)	Domestic Equity (T=1)	Domestic Equity (T=1.5)
Logical	89%	87%	88%
Unclear	11%	10%	5%
Unrelated	0%	3%	7%
	Sector (T=0)	Sector (T=1)	Sector (T=1.5)
Logical	95%	92%	90%
Unclear	4%	6%	4%
Unrelated	1%	2%	6%
ChatGPT 4.0			
	Domestic Equity (T=0)	Domestic Equity (T=1)	Domestic Equity (T=1.5)
Logical	82%	78%	79%
Unclear	15%	20%	16%
Unrelated	3%	2%	5%
	Sector (T=0)	Sector (T=1)	Sector (T=1.5)

Panel A: Financial Analyst			
ChatGPT 3.5			
	Domestic Equity (T=0)	Domestic Equity (T=1)	Domestic Equity (T=1.5)
Logical	85%	89%	85%
Unclear	13%	2%	11%
Unrelated	2%	9%	4%
	Sector (T=0)	Sector (T=1)	Sector (T=1.5)
Logical	97%	95%	94%
Unclear	1%	3%	5%
Unrelated	2%	2%	1%
ChatGPT 4.0			
	Domestic Equity (T=0)	Domestic Equity (T=1)	Domestic Equity (T=1.5)
Logical	79%	83%	75%
Unclear	17%	12%	17%
Unrelated	4%	5%	8%
	Sector (T=0)	Sector (T=1)	Sector (T=1.5)
Logical	89%	90%	85%

- **Nhận xét:**

- **ChatGPT 3.5 hoạt động tốt hơn ChatGPT 4.0 ở cả hai thị trường. Khuyến nghị đầu tư nhóm ngành có độ chính xác cao hơn đầu tư vào các công ty cụ thể, đặc biệt ở Trung Quốc, có thể do chính sách Nhà nước có tính định hướng rõ ràng hơn tin tức từ Wall Street Journal.**
- **Hệ số “logical” cao thể hiện khả năng nhận diện và hiểu sâu của ChatGPT**
- **ChatGPT có thể nhận diện cơ hội tốt hơn thể hiện qua các thông số (Logical, Unclear, Unrelated)**

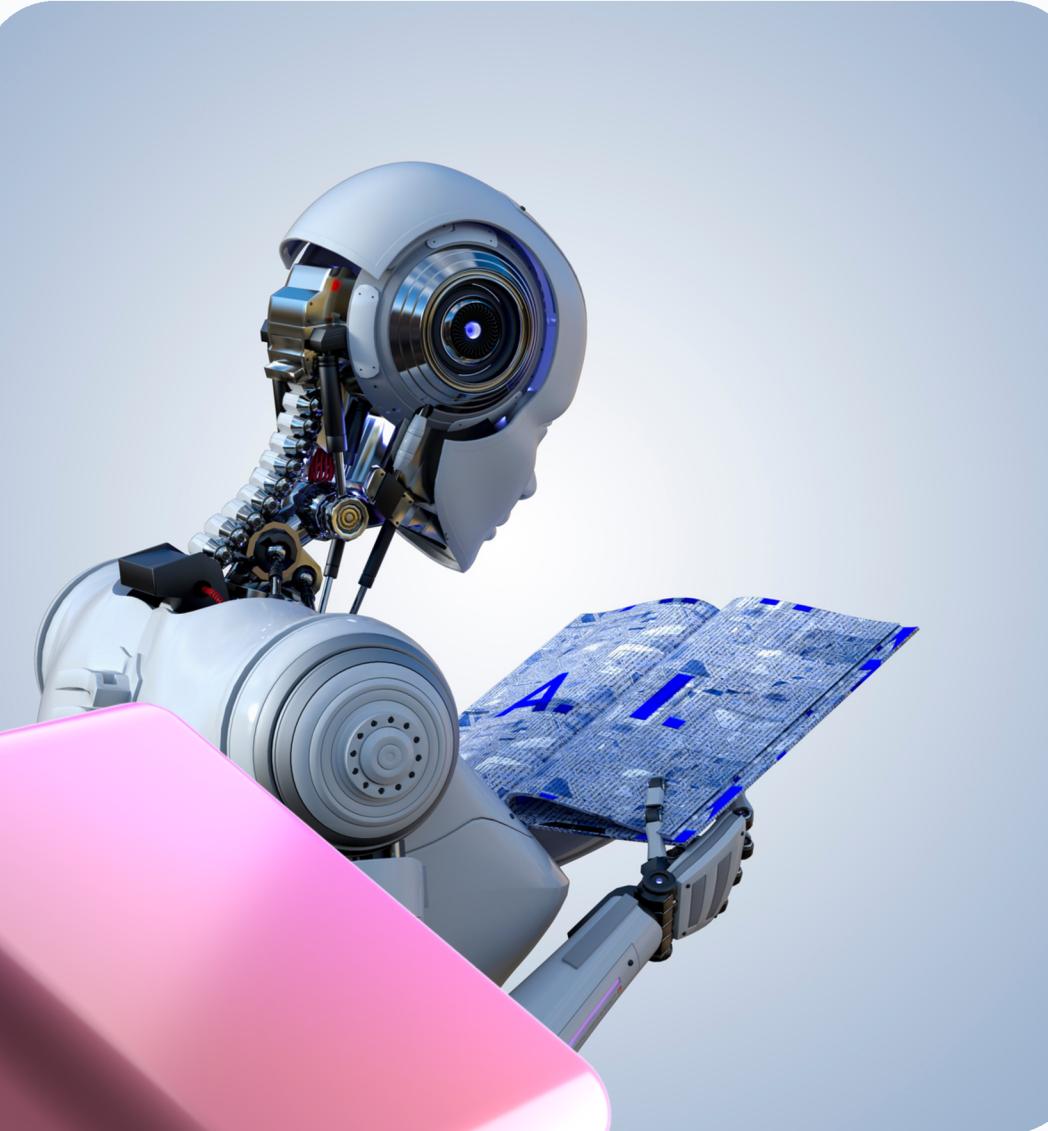
# HẠN CHẾ CỦA CHATGPT

- Phụ thuộc vào dữ liệu đầu vào
- Không có khả năng tính toán tài chính phức tạp
- Overfitting vào dữ liệu quá khứ
- ChatGPT có thể tạo danh mục rủi ro cao
- ChatGPT chưa được tối ưu cho tài chính
- Cần kiểm thử trên nhiều dữ liệu khác nhau

⇒ Hướng phát triển: Fine-tune mô hình, kết hợp AI + dữ liệu thị trường



# TIỀM NĂNG ỨNG DỤNG CỦA CHATGPT TRONG ĐẦU TƯ



**Hỗ trợ nhưng không thay thế**

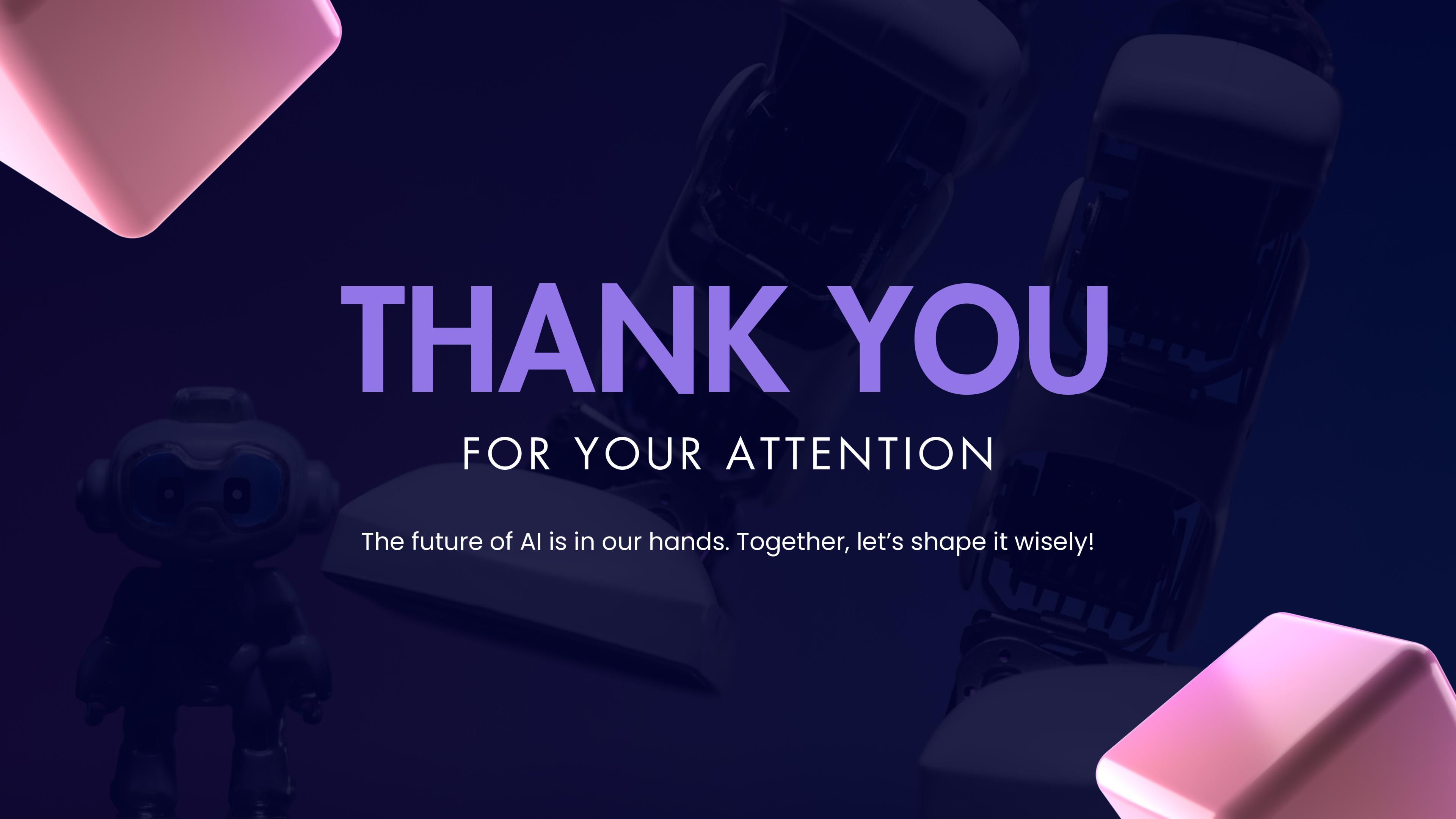
ChatGPT hỗ trợ quyết định, nhưng không thay thế hoàn toàn con người

**Tối ưu danh mục, thích ứng dữ liệu**

Khả năng xây dựng danh mục đầu tư hiệu quả hơn, thích ứng với nhiều nguồn dữ liệu khác nhau

# TÓM TẮT KẾT LUẬN

- 1 ChatGPT có thể tạo danh mục đầu tư hiệu quả
- 2 ChatGPT vượt trội hơn so với các phương pháp truyền thống
- 3 ChatGPT có thể tạo danh mục với three-factor alpha lên đến 3%/tháng, đặc biệt trong môi trường thị trường bị ảnh hưởng bởi chính sách.
- 4 Hiệu suất danh mục phụ thuộc vào temperature: Temperature cao cho kết quả đa dạng hơn, nhưng có thể đưa ra các lựa chọn không nhất quán hoặc kém chính xác.



# THANK YOU FOR YOUR ATTENTION

The future of AI is in our hands. Together, let's shape it wisely!