

T. Rowe Price

Human v.s. AI in Financial Analysis

*Cathy Shan, Ahmed Wakrim, Maya Walcher, Nicholas Wong,
David Xiang, Iris Xiao, Jinghan Xu, Rei Yamahara*

MIT SLOAN SCHOOL
OF MANAGEMENT



- 1. Introduction & Definition**
- 2. Literature Review**
- 3. Empirical Test on GPT For Equity Report**
- 4. Model Evaluation**
- 5. Conclusion**

AGENDA



FINANCE

JPMorgan Chase is giving its employees an AI assistant powered by ChatGPT maker OpenAI

PUBLISHED FRI, AUG 9 2024 7:00 AM EDT UPDATED FRI, AUG 9 2024 10:46 AM EDT



Hugh Son
@HUGH_SON

WATCH LIVE

KEY POINTS

- GenAI is transforming finance operations
- GenAI's potential in decision-making

Recent News about GenAI in Finance...



THE WALL STREET JOURNAL.

CIO JOURNAL

Six Months, Thousands of GPTs and Some Big Unknowns: Inside OpenAI's Deal With BBVA

The bank is evaluating its next steps after seeing early productivity gains from OpenAI's enterprise tool

By *Isabelle Bousquette* [Follow](#)

Nov. 21, 2024 7:00 am ET

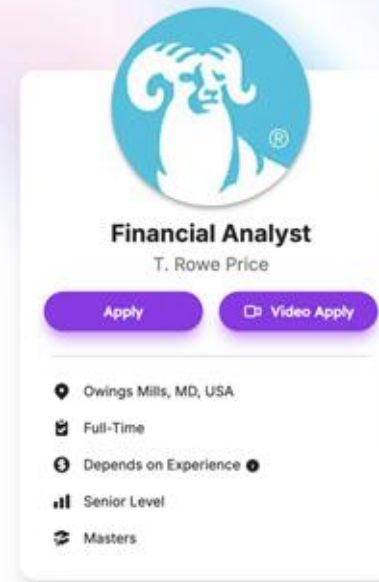
What is a Financial Analyst?

- Fundamental Research
- Stocks Recommendation
- Communication

Research Analyst – Asia Equities

We are looking for someone to join us as a Research Analyst covering stocks in Asia to:

- Conduct fundamental research on stocks in line with our investment framework.
- Build and maintain financial models that help in our valuation of the companies.
- Assess the quality of the companies covered using our proprietary assessment.
- Support your investment research with boots-on-the-ground research activities where you will meet companies, industry experts, regulators, customers, and suppliers.
- Provide a high-conviction top pick list to the EM strategy team and EM portfolio managers to be added to our client portfolios.
- Assume responsibility for the performance of your recommended stocks.



Responsibilities:

1. Analyze financial data and develop reports to support business decisions.
2. Develop financial models for forecasting and budgeting.
3. Collaborate with other departments to ensure that financial analysis is accurate and up-to-date.
4. Monitor industry trends and financial markets to identify potential opportunities and risks.
5. Prepare presentations on financial results and analysis for management.
6. Participate in financial planning and forecasting activities.
7. Develop and maintain effective relationships with internal and external stakeholders.
8. Assist in the preparation of audit documents.
9. Ensure compliance with laws and regulations related to financial reporting.
10. Provide guidance and mentorship to team members.

Equity Research Analyst New York, NY

1. Analyze the financial and operational performance of industry groups or individual companies and conduct proprietary investment research and fundamental financial analysis.
2. Develop financial models to project the future performance of companies under coverage. Perform company valuation based on company financials. Examine financial reports, organize data, and collect materials to present and summarize investment ideas.
3. Build financial models to forecast sales of key products under different assumptions.
4. Write research reports for initiations of coverage, analysis of financial data, and industry/company news.
5. Delve into intellectual properties, the patent law, and litigations related to biotech and pharmaceuticals.
6. Conceive and execute differentiated project ideas. Interface with company management teams, internal sales and trading personnel, and institutional investors.
7. Write research reports based on fundamental financial analysis, evaluation of company earnings release, and guidance.
8. May telecommute.

Workflow of a Financial Analyst



Data Collection



- **Public Data:** SEC filings, industry reports, market data
- **Private Data:** Executive insights, proprietary datasets, expert calls

Research



- Analyze financial data to evaluate company performance
- Build forecasting models to predict future trends
- Create detailed and actionable research reports

Recommendation



- Provide buy, sell, or hold recommendations based on research
- Other key insights

Communication



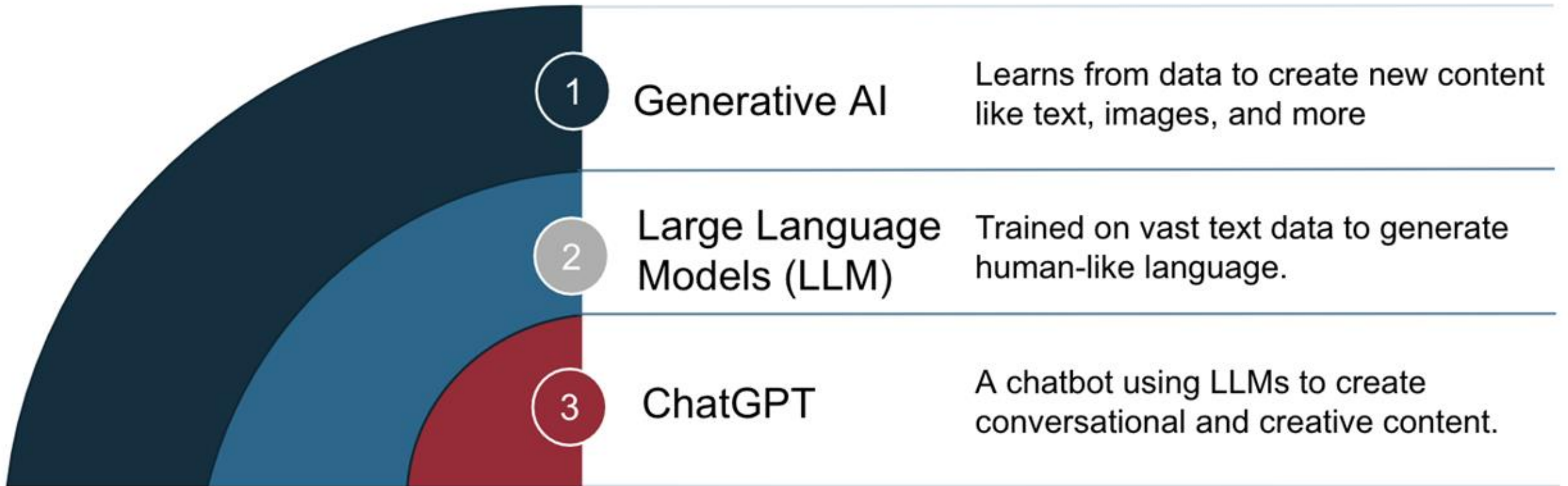
- Engage with clients and stakeholders to share insights
- Collaborate with internal teams and external companies

What is GenAI

Content Creation

Data Understanding

Applications





So...

Can GPT replace Financial Analysts?

NOT YET!

- 1. Introduction & Definition**
- 2. Literature Review**
- 3. Empirical Test on GPT For Equity Report**
- 4. Model Evaluation**
- 5. Conclusion**

AGENDA



Evaluation of LLMs on mock CFA Exams

Table: LLMs' ability to pass each CFA level

Provider	Model	Level I		Level II		Level III	
		L	U	L	U	L	U
OpenAI	GPT-3.5 Turbo	✓	✗	✗	✗	✗	✗
	GPT-4 Turbo	✓	✓	✓	✓	✗	✗
	GPT-4o	✓	✓	✓	✓	✗	✗
Anthropic	Claude 3 Opus	✓	✓	✓	✓	✗	✗
Mistral	Mixtral-8x7B	✓	✗	✗	✗	✗	✗
	Mixtral-8x22B	✓	✗	✓	✗	✗	✗
	Mistral Large	✓	✗	✓	✗	✗	✗
Google	Gemma 2B	✗	✗	✗	✗	✗	✗
	Gemma 7B	✗	✗	✗	✗	✗	✗
Meta	LLaMA 3 8B	✗	✗	✗	✗	✗	✗
	LLaMA 3 70B	✓	✗	✗	✗	✗	✗
	LLaMA 3 8B + RAG	✓	✗	✗	✗	✗	✗
	LLaMA 3 70B + RAG	✓	✓	✓	✗	✗	✗
Cohere	Command R+	✗	✗	✗	✗	✗	✗
Microsoft	Phi-3-mini	✓	✗	✗	✗	✗	✗
Ai2	OLMo 7B	✗	✗	✗	✗	✗	✗

Some LLM can pass Level I and Level II

No LLM can pass Level III

Note: Lower bound score L ($\geq 60\%$) and upper bound score U ($\geq 70\%$)

References: The State of the Art of Large Language Models on Chartered Financial Analyst Exams. Mahfouz et.al (2024)

GPT demonstrates effectiveness in processing information



[...] Q4 was another strong period for Parks, Experiences and Products, which continue to deliver phenomenal results despite the impact of Hurricane Ian. I want to thank the amazing cast members in Orlando who went above and beyond to help keep our guests safe and entertained during the storm. From protecting the many animals at Disney Animal Kingdom to packing thousands of meal kits, to donating and delivering emergency supplies to the community, I am so proud of how our team came together to support our guests, our neighbors and each other.

Our Parks team is laser focused on enhancing the guest experience in creating those magical memories I mentioned at the outset. This focus, along with the investments we made to bringing new attractions and experiences to our domestic parks, is generating consistently strong demand, which, on many days, exceeds our current capacity. And we continue to manage attendance levels with a focus on providing guests with the highest quality experience in enhancing our Park's overall financial performance.

One of the things our guests love most is the opportunity to celebrate at our parks, as evidenced by the post-pandemic return and sell-out of special ticketed events like Boogie Boogie Bash and Mickey's Not So Scary Halloween party. I visited Disneyland with my family just before Halloween and the celebration was phenomenal. Tickets for Mickey's Very Merry Christmas party at Walt Disney World has now officially gone on sale and over half of all dates have already sold out.

As you know, we are about to embark on the company's 100th anniversary celebration. The fun kicks off in our parks on January 27 at Disneyland, where we'll unveil new platinum infused decor, Premier II, all new nighttime spectaculars and opened a highly anticipated Mickey & Minnie's Runaway Railway attraction. This is only one part of what will be the largest cross-company celebration in Disney's history with activations around the world, and we're so excited for fans and families to join us.

At our international parks, Disneyland Paris is enjoying a great resurgence. Our fantastic new Marvel Avengers Campus opened on July 20, and guests love a highly immersive and dynamic environment of the first-ever Marvel-themed land in Europe. Prior to the recent closure of Shanghai Disney Resort, we are seeing positive momentum there and at Hong Kong Disneyland. We are hopeful that the situation will improve and are thinking of all of our employees there as we manage through the challenging COVID environment. Our Disney Cruise line is showing strong signs of recovery. The new Disney Wish is in high demand, and we've seen a ramp-up in bookings for our base fleet. [...]

[...] Despite the challenges posed by Hurricane Ian, the Parks, Experiences, and Products segment exhibited resilience. The Orlando team demonstrated remarkable dedication by ensuring guest safety, protecting animals, and providing community assistance during the storm. Their focus on enhancing guest experiences remained a top priority, leading to investments in new attractions and experiences, even as they managed overwhelming demand. Celebratory events also gained immense popularity. Disney's international ventures also thrived. Disneyland Paris benefited from the Marvel Avengers Campus, and the Asian parks performed well before unexpected closures. Disney Cruise Line, especially the new Disney Wish, experienced high demand. [...]

- Demonstrated economic utility of GPT summaries
- Highly effective in summarizing complex financial disclosures

~80%

Reduction in document length

GPT in Financial Analysis

Three Important Papers:

- **Earnings Changes Prediction**

*“Financial Statement Analysis with Large Language Models”
(Kim et al. 2024)*

Using financial statements to predict direction of future earnings.

- **Equity Stock Rating**

*“AI in Investment Analysis: LLMs for Equity Stock Ratings”
(Papasotiriou et al. 2024)*

Using a combination of financial fundamentals, news and historical stock returns to predict quantile of stock returns (ratings).

- **Stock Selection**

“Can Large Language Models Beat Wall Street? Unveiling the Potential of AI in Stock Selection” (Fatouros et al., 2024)

Using a combination of financial fundamentals, news, historical stock returns, and macroeconomic data to generate buy/hold/sell signals and design trading strategies.



Insight 1: Importance of Chain-of-Thought (CoT) Prompting



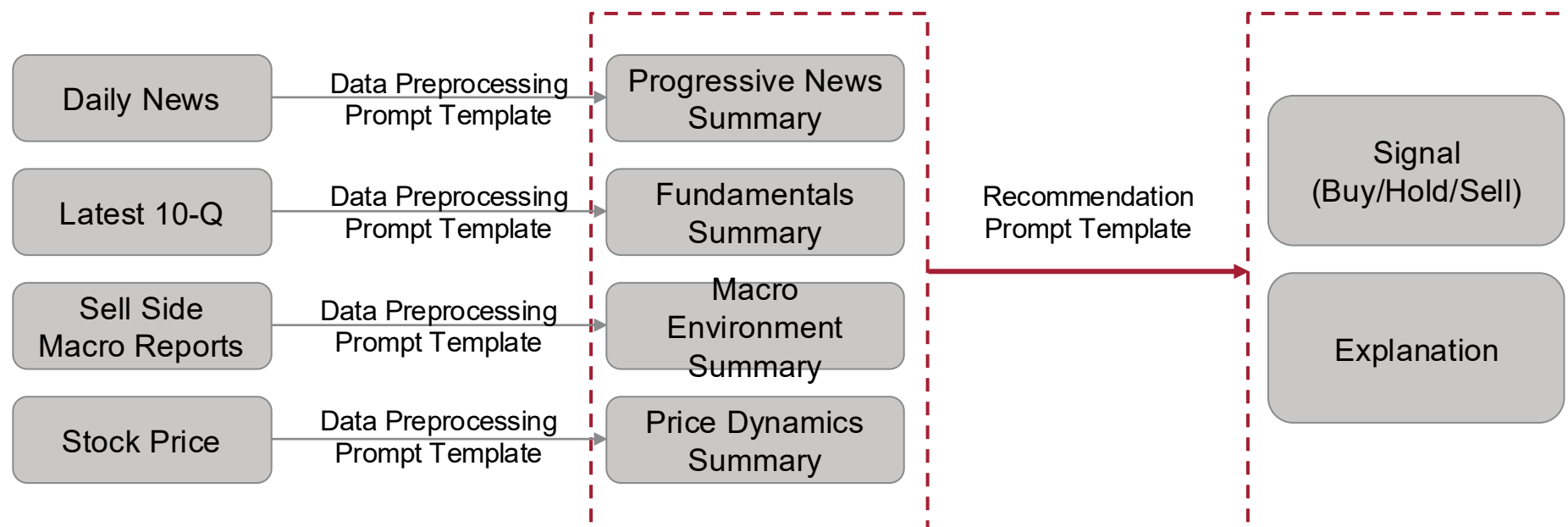
Earnings Changes Prediction (Kim et al., 2024)

- Chain-of-Thoughts (CoT) Prompt: identify changes, compute financial ratios, interpret ratios and synthesize insights.
- Results: CoT reasoning improved GPT-4's prediction accuracy from 52.71% (simple prompts) to 60.31%.

Stock Selection (Fatouros et al., 2024)

Highlights the role of CoT prompting in analyzing diverse data sources and emphasized step-by-step analysis.

Figure: Pipeline of MarketSenseAI



Insight 2: GPT Outperform Human Analysts and the Market



Earnings Changes Prediction (Kim et al., 2024)

Earnings Changes Prediction (Kim et al., 2024):

GPT-4's accuracy (60.31%) in **earnings change directions** prediction outperformed human analysts (52.71%), and matches specialized Artificial Neural Network (ANN) models (60.45%).

Stock Selection (Fatouros et al., 2024)

MarketSenseAI (Fatouros et al., 2024)

Market cap strategy powered by GPT-4 delivered cumulative returns up to 66.22%, outperforming the S&P 100 by 23%.

Table: Performance (from December 2022 to March 2024)

Strategy	Total Return (%)	Sharpe	Sortino	Vol (%)	Win Rate (%)	Max Ddn (%)
S&P100	43.27 (43.20)	2.86	4.61	16.17	N/A	9.24
Signal-Cap-L	66.22 (65.25)	2.90	4.95	22.81	65.88	9.64
S&P100-Eq	25.22 (25.12)	1.98	3.02	14.72	76.36	10.66
Signal-Eq	35.48 (32.94)	2.49	3.87	15.76	65.68	8.47
Signal-Eq-L	35.79 (34.82)	2.41	3.75	16.44	65.02	9.00

Note: Values in bold indicate the best scores among the equally weighted (Eq) or market-cap weighted (Cap) strategies. "Signal" represents strategies utilizing buy/hold/sell signals, while "L" denotes long-only strategies. Values in parentheses represent total returns after transaction costs (5bps/trade).

Insight 3: Complementarity Between GPT and Human Analysts

Earnings Changes Prediction (Kim et al., 2024); Equity Stock Rating (Papasotiriou et al., 2024)



GPT and human analysts have distinct strengths:

- Human Analysts:
 - Better at predicting challenging companies, such as small or loss-making companies, where contextual and domain-specific insights are crucial.
 - Human analysts outperform in long-term predictions (18 months).
- GPT:
 - Adds value where analysts show bias or disagreement.
 - GPT excels in short-term (1–3 months) and medium-term (3–12 months) predictions.

If we combine the two, they give us complementary predictions:

- Regression analysis demonstrates GPT and analysts distinctly contribute to earnings forecasts.

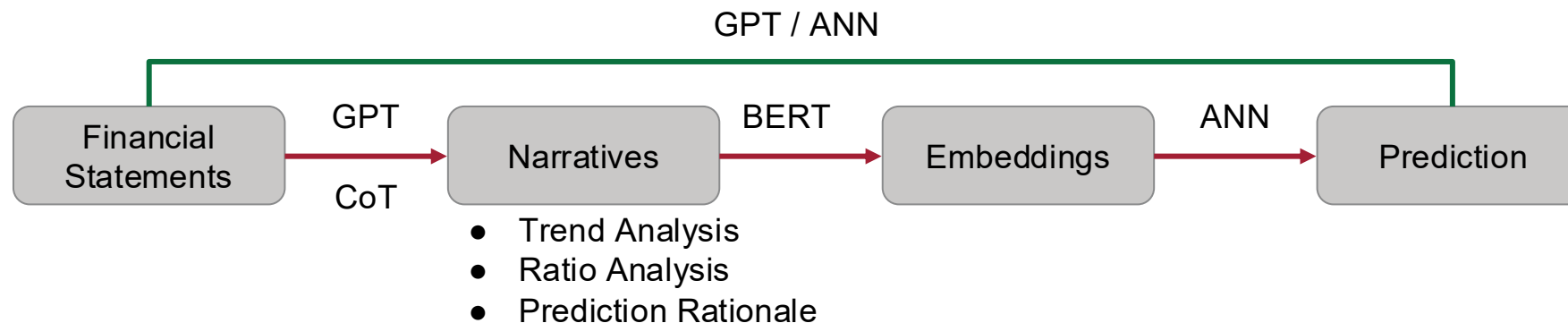
Insight 4: Explanation and Informative Value of GPT Outputs



Earnings Changes Prediction (Kim et al., 2024)

Earnings Changes Prediction (Kim et al., 2024).

A key strength of GPT: its ability to not just make accurate predictions but also **explain its reasoning** through narratives.



Results

- Results are **comparable** to direct GPT and ANN predictions, achieving 59% accuracy and a 65% F1-score, which suggests that GPT-4's narrative outputs encode informative insights about financial performance, and served as a basis for GPT's predictive ability.

What These Papers Addressed and Missed

Breadth v.s. Depth



Showcasing Strength of GPT in Breadth

The strengths of GPT in finance showcased in these papers:

- GPT excels at **analyzing a wide range of companies at scale**.
- It efficiently **summarizes data** and **synthesize information** across multiple dimensions for predictions.


Limited Results Evaluating Depth

- **Result are limited to high-level outcomes:**

Predictions focus on high-level outcomes, like earnings changes or stock ratings, **lacking the granularity** required for detailed financial decision-making.

- **Didn't explore GPT's ability of deep reasoning:**

Despite the papers demonstrating that GPT surpasses human analysts in accuracy and provides valuable explanations, they did not evaluate GPT's capability for **deep reasoning**.

A large, light blue, stylized quotation mark graphic on the left side of the slide.

Now our team wants to explore:
Can GPT Match Analyst-Level Depth?

- 1. Introduction & Definition**
- 2. Literature Review**
- 3. Empirical Test on GPT For Equity Report**
- 4. Model Evaluation**
- 5. Conclusion**

AGENDA



Testing Framework: Layered Model for GenAI Performance Evaluation

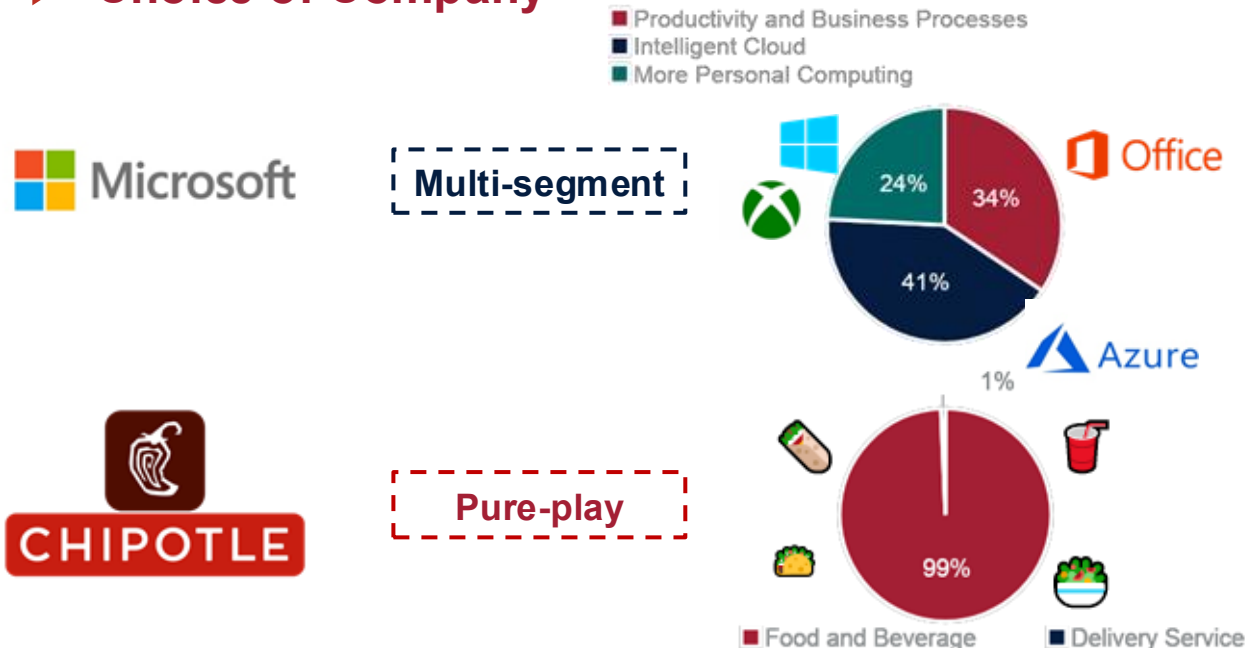


► Goal for ChatGPT 4o

Mimic the role of a sell-side equity research analyst

- Draft an equity research report
- Base on the stock's most recent quarterly earnings release
- Provide investment thesis & recommendation for a stock

► Choice of Company

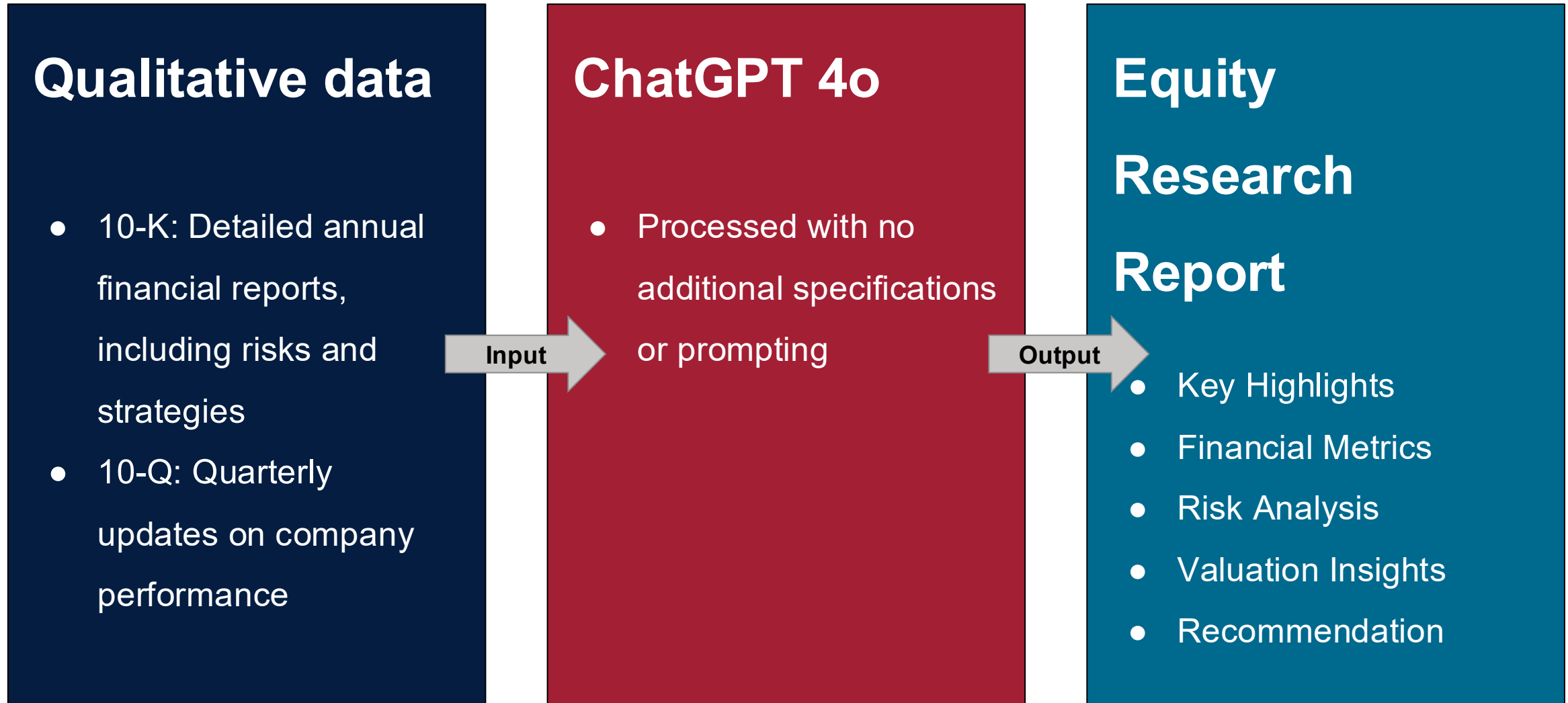


► Model Structure

Human participation

	Levels of Input	Prompt Complexity
Layer 1	low	low
Layer 2	high	medium
Layer 3	high	high

Level 1: Minimum Inputs & No Prompt



Methodology & Purpose

Assess AI's ability to extract insights from raw regulatory filings.

No external data or analyst insights were used.

Prompt

"I want you to build a sell-side research report on Microsoft. I will give you some qualitative data and I want you to write a sell side analyst report with price target and sell/neutral/buy recommendation. I will also give you a template (which is a bofa sell side analyst report for lyft) for you to see the format I am looking for for this analysis."



Input



MSFT_2025_10Q.pdf
PDF



bofa_lyft_report.pdf
PDF



Reports.pdf
PDF



MSFT_10K.pdf
PDF



MSFT earnings call.docx
Document

Level 2: Enhanced Model with High-Level Inputs



Inputs

- 1 Financial Report:** 10-K, 10-Q, Earnings Call Transcripts from the past 1 year
- 2 Industry Report:**
 - Market Sizing
 - Competitive Landscape Analysis
- 3 Historical Analyst Report:**
 - Analyst perspectives on performance, valuation and rating from the past 6 months
- 4 Gross Stock Return**
 - Historical Stock Return
 - Benchmark Index Return for rating evaluation

22

+

Prompt

- 1 Extraction Guidance:**
 - Financial/Operating Data
 - Calculating Key Ratios
 - Contextual Insights

'Based on Transcripts context, summarize management outlook & strategy with citing'

- 2 Production Guidance:**
 - Rating Rule:** refer to Morgan Stanley
 - Predict next quarter stock return

Format:

Exceed benchmark return	-> BUY
Near benchmark return.	-> HOLD
Below benchmark return	-> SELL

Performance Update	Key Catalysts & Drivers
Rating & Risk	Valuation

Management Outlook and Strategy

Area	Management Direction	Future Forecast
AI Leadership	"Infusing AI across every layer"; large-scale AI deployments (TranscriptFY24Q2) (TranscriptFY24Q3)	Continued innovation; anticipated leadership in AI-based cloud and enterprise applications (TranscriptFY24Q3)
Cloud Dominance	Expanding Azure services and deepening AI integration (TranscriptFY24Q3)	Billion-dollar multi-year commitments; focus on retaining enterprise customers (TranscriptFY24Q3)
Market Expansion	Strategic acquisitions (Activision); new hardware (AI PCs) (TranscriptFY24Q3)	Broader content portfolio; new AI-driven consumer experiences (TranscriptFY24Q3)
Operational Efficiency	Prioritizing cost control alongside strategic growth investments (TranscriptFY24Q1) (TranscriptFY24Q3)	Operating margins expected to improve; increased focus on higher-margin businesses like E5 subscriptions (TranscriptFY24Q3)

Methodology of GPT's Choice



Benchmark Index



MSCI USA IMI Information Technology Index



Dow Jones U.S. Restaurant & Bars Index

Rationale Given

- Sector-specific Relevance
- Similar Market Risks: reflect broader trends
- Presence in the index: offer direct comparability

TOP 10 CONSTITUENTS

	Float Adj Mkt Cap (USD Billions)	Index Wt. (%)
APPLE	3,290.92	21.25
NVIDIA	3,265.90	21.09
MICROSOFT CORP	2,869.11	18.52

Stock Return Prediction Model

Stock return prediction in all attempts (6/6) comes from linear regression of the company's historical return, but against different data:

Company Level Historical Revenue Growth 2

Industry Level Overall Market Growth 1


Macro Level Consumer Consumption Change 3




Inconsistency Issue

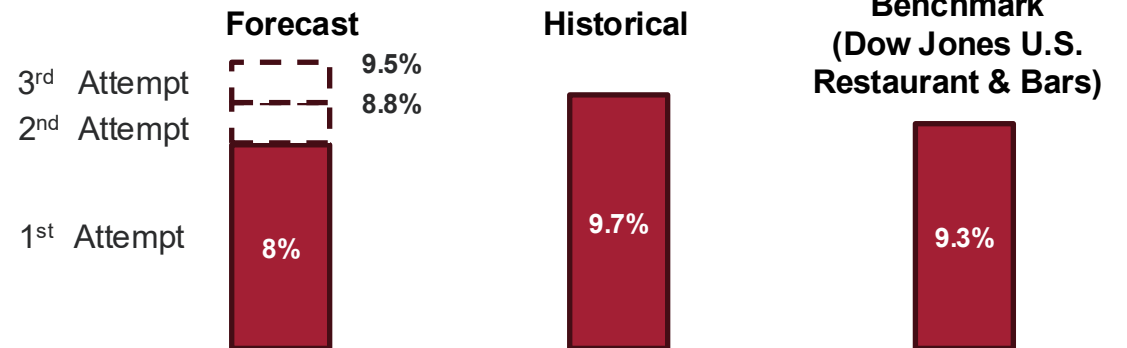
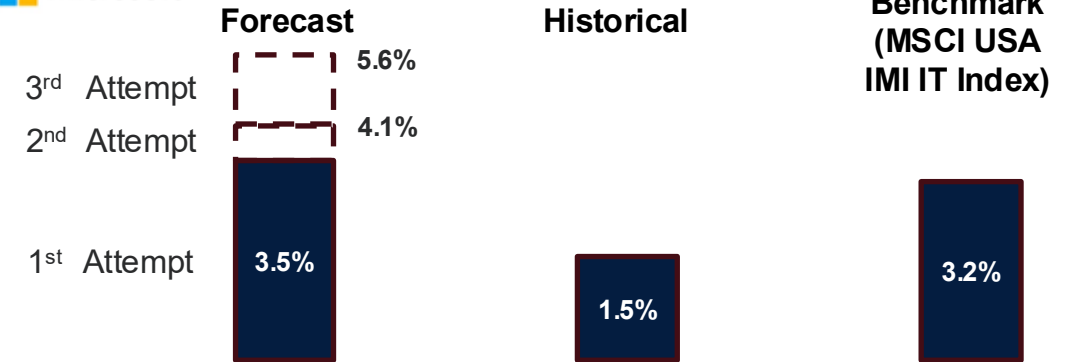
Performance Comparison

➤ Against Analyst Consensus

 Microsoft	Analysts Per level	GPT Per level	
		Layer 1	Layer 2
	01-Nov-2024	01-Dec-2024	
1 - Buy	54	5	3
2 - Hold	4	0	0
3 - Sell	0	0	0
Rec Mean	1.07	1.00	1.00

 CHIPOTLE	Analysts Per level	GPT Per level	
		Layer 1	Layer 2
	01-Nov-2024	01-Dec-2024	
1 - Buy	25	5	2
2 - Hold	10	0	1
3 - Sell	0	0	0
Rec Mean	1.29	1.00	1.33

➤ Against Past Quarter Stock Return (Ending Oct.31st, 2024)



Layer 1 & 2 Delivery and Limitation



Delivery

Financial Summary

1. Revenue Growth:

- Total: \$65.6B (+16% YoY)
- Intelligent Cloud: \$23.9B (+19% YoY)
- Productivity & Business Processes: \$18.6B (+13% YoY)

2. Profitability:

- Gross Margin: \$45.5B (69.3%)
- Operating Income: \$30.6B (+13.6% YoY)

Simple Valuation Projection

Quarter	Revenue (\$B)
Q2 FY25	68
Q3 FY25	70.5
Q4 FY25	72.8
Q1 FY26	75.2

...And Offer Rationale

- Strong AI adoption drives cloud growth
- Activision integration enhances gaming
- Operating leverage offsets margin pressures

Limitation/Challenge

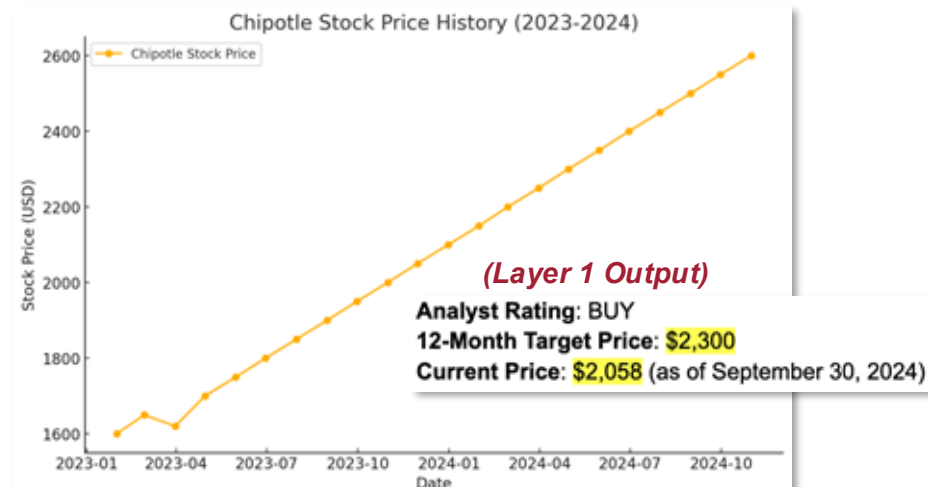
Low Information Density

- Lack analytical depth
- Lack forward-looking guidance
- Lack logical reasoning to connect facts & statements

→ Prompt with Expertise

Inaccuracy & Inconsistency

- Real Price: 60.76 (Oct, 2024)



→ Regulation On Sourcing

Level 3 - Custom GPT: Initialization of model



Custom GPT

- A tailored AI tool where users can customize its behaviour and knowledge base
- Provides accurate, industry-specific outputs

Initialising GPT by inserting role description

“Financial Analyst is a Bank of America Financial Equity Analyst that creates research reports for Chipotle, a company in the fast-food restaurant industry. Your analysis should include financial modeling, valuation and a written investment recommendation. You will be provided with an exact format and example template to follow, which will be provided at the very end.

Use only the attached qualitative and quantitative data for your analysis. Do not use additional data from external websites.

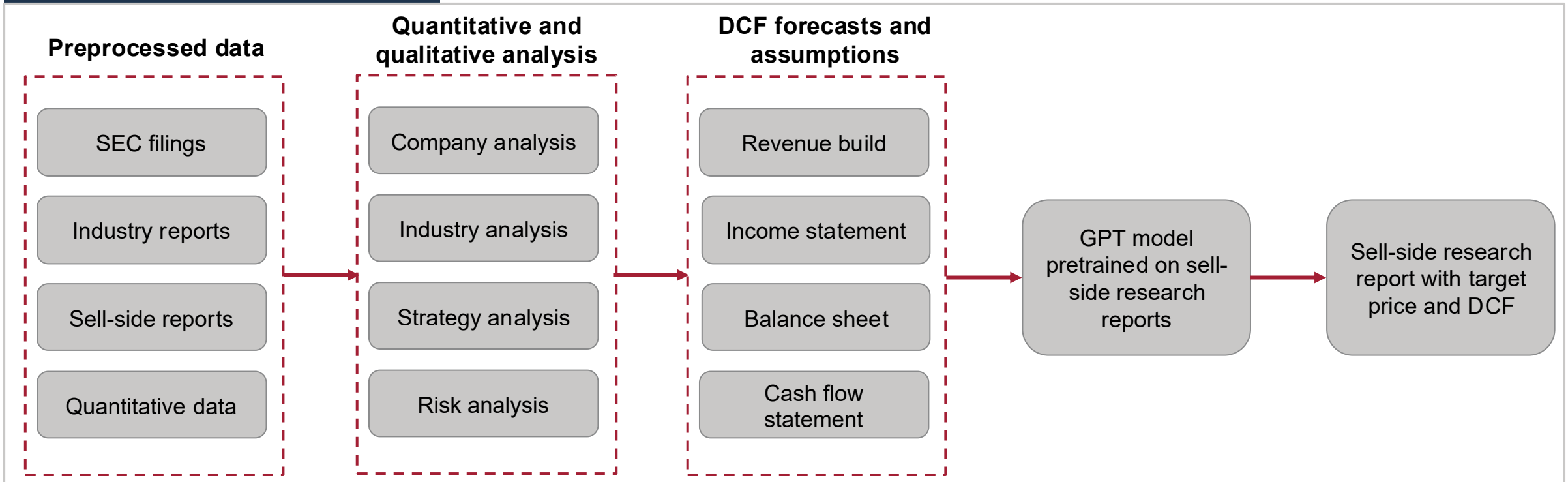
Modular Prompting Framework



Modular Prompting

Breaking complex tasks into smaller, focused prompts for improved accuracy and coherence

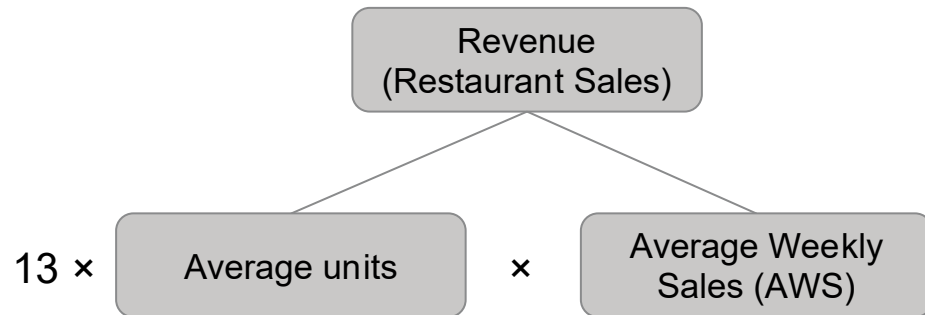
Model Framework



Prompt Example: Revenue Build



Revenue Breakdown



- **Chipotle AWS** = Chipotle AWS (same quarter previous year) × (1 + YoY AWS growth rate)
- **YoY AWS growth rate** = Same-store sales (SSS) + “AWS - SSS Gap”
- **Same-store sales (SSS)** = Pricing + Mix + Transaction
- **Average units** = average units of current and previous quarter units

Specific example: prompt for forecasting Chipotle units opened/closed in quarter

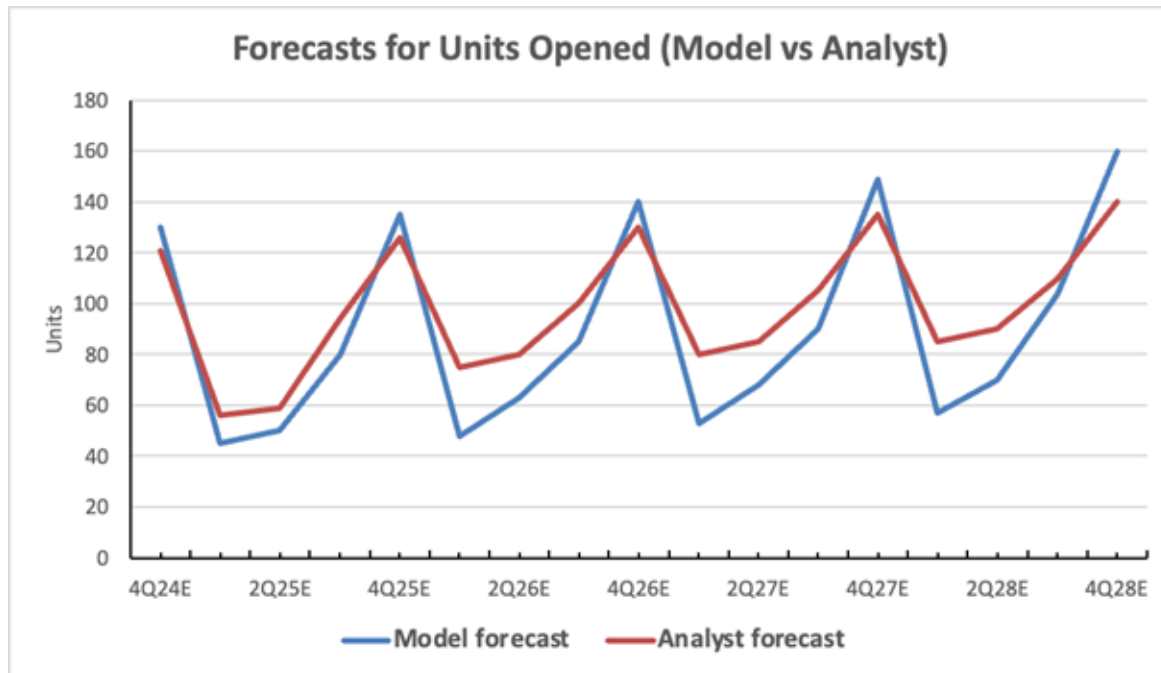
“Based on company analysis, industry trends, forward guidance, management targets, macroeconomic factors, and historical data on store openings and closures, estimate the future trend for the number of Chipotle units opening and closing over the next 17 quarters (from 4Q2024E to 4Q2028E).”

1. *Provide your estimates in an Excel spreadsheet, clearly organizing the data by quarter and including any relevant visualizations (e.g., trendlines or charts) to support your forecast.*
2. *Write a detailed explanation of your methodology, outlining the step-by-step reasoning behind your estimates and referencing specific quantitative data and trends (including insights from previous prompt answers).*
3. *Assess the reasonableness of your estimates by comparing them with historical trends, company targets, industry benchmarks, and macroeconomic conditions. Discuss how you incorporated previous prompt answers into your analysis and address any potential sources of variability (e.g., economic downturns, shifts in consumer demand, or operational challenges).*

Prompt Example: Revenue Build



Comparing forecasted units to MS forecasted units



Shortened reasoning for forecasts by GPT model

“Growth Aligned with Trends: The forecast reflects Chipotle’s long-term goal of 7,000 units, scaling historical growth of 50 units per quarter and 2.6 YoY unit growth to account for accelerated expansion via Chipotlanes and international markets.

Realistic Growth: Projects unit opening follows a balanced YoY growth, scaling early growth and moderating later as the company matures.

Seasonality & Feasibility: Unit openings are aligned with Chipotle’s seasonal trends and operational strategy. Consistent with historical data, closures remain stable at around 3 units per quarter, reflecting disciplined growth.”

Provide Prompt and template

Pre-training the model

1. Input BofA sell-side research reports
2. Extract and replicate the formatting and style of BofA reports
3. Generate a standardized research report template

Final Prompt

“You are a sell-side equity research analyst at Bank of America tasked with creating a sell-side research report for Chipotle. Using the data from Chipotle’s Q4 2024 report, create an outlook for Q1 2025 using the template provided and fill out the quantitative metrics.”

Lyft, Inc.

Turnaround taking time, but rider signals still trending positively; Maintain Buy

Reiterate Rating: BUY | PO: 16.00 USD | Price: 9.67 USD

Mixed quarter with bookings miss, revenue/EBITDA beat

Reported bookings/rev/EBITDA of \$4.02bn/\$1.44bn/\$105mn, which was 2% below/4% above/6% above Street at \$4.07bn/\$1.39bn/\$99mn. Rides grew 15% y/y, users grew 10%, and Bookings grew 17% y/y (4pt decel on a 2pt easier comp), and came toward the low end of guidance at \$4.0-\$4.1bn. A positive was the strong take rate at 35.7%, up 1pt Q/Q, reflecting healthy supply with record driver sign ups in 2Q, and fastest pickup times in 4 years. EBITDA at \$105mn was a modest beat, while FCF at \$276mn was well above but benefitted from insurance accrual timing. Possibly biggest surprise was Prime Time reduced by 25% Q/Q, positive for long-term platform health but N.T. profit headwind as surge- rides have historically been accretive to EBITDA vs. standard rides.

3Q outlook suggests impact from less Prime Time

3Q outlook for bookings of \$4.0bn-\$4.1bn was below Street at \$4.14bn while EBITDA outlook of \$90mn-\$95mn also missed at \$104mn prior. Mgmt. maintained FY'24 commentary, and noted “good visibility” into insurance renewal with increase coming in below LY’s 17% jump. Mgmt also announced Price Lock, an effective price cap for those that sign up for rides in advance, though it seems rollout will take several quarters.

Estimate changes: raising revenue, lowering EBITDA

For 3Q, we lower bookings 2%, raise revenue 7% on Take Rate efficiency, but lower EBITDA 2%. For 2025E, we are at Revenue/EBITDA of \$6.54bn/\$528mn vs. \$6.43bn/\$584mn prior. Take rate & Media can support EBITDA growth from here, while unit economics headwinds include insurance, less Prime Time, and mix-shift. Lyft is also investing in rider incentives, with S&M up 79% Y/Y in 2Q (and we project Q/Q growth).

Growth trailing Uber, but inexpensive valuation; Buy

Lyft’s mix-shift to lower-priced rides via Wait-n-Save/Price Lock while also reducing Prime Time is likely limiting EBITDA upside, but supports long-term health as the secondary US rideshare platform. With progress in headwinds like insurance, SBC, and potential for AV partnerships, Lyft’s valuation can converge closer to Uber’s longer-term. Trading at just 0.4x our 2025E revenue & 5x EBITDA, we still like the risk/reward potential from here (Uber at 3x 2025E revenue & 18x 2025E EBITDA). We lower our PO to \$16 (vs. \$20 prior) based on our DCF that now assumes slightly lower growth & margin estimates, as well as a 10x terminal multiple vs. 12x prior; Reiterate Buy rating.

Estimates (Dec) (US\$)	2022A	2023A	2024E	2025E	2026E
EPS	(1.48)	0.65	0.90	1.25	1.80
GAAP EPS	(4.47)	(0.88)	0.01	0.23	0.68
EPS Change (YoY)	-131.3%	NM	38.5%	38.9%	44.0%
Consensus EPS (Bloomberg)			0.70	0.90	1.01
DPS	0	0	0	0	0
Valuation (Dec)					
P/E	NM	14.9x	10.7x	7.7x	5.4x
GAAP P/E	NM	NM	967.0x	42.0x	14.2x
EV / EBITDA*	NM	16.2x	9.9x	6.8x	4.6x
Free Cash Flow Yield*	-10.6%	-7.5%	13.6%	12.5%	20.4%

* For full definitions of *Qinetiq* measures, see page 6.

07 August 2024

Equity

Key Changes

(US\$)	Previous	Current
Price Obj.	20.00	16.00
2024E Rev (m)	5,439.4	5,687.0
2025E Rev (m)	6,430.0	6,536.5
2026E Rev (m)	7,477.7	7,385.5
2024E EPS	0.93	0.90
2025E EPS	1.39	1.25
2026E EPS	1.85	1.80

Michael McGovern
Research Analyst
BofA
+1 415 676 3520
mmcgvorn@bofa.com

Stock Data

Price	9.67 USD
Price Objective	16.00 USD
Date Established	7-Aug-2024
Investment Opinion	C-1-9
52-Week Range	8.85 USD - 20.82 USD
Mkt Val (mn) / Shares Out (mn)	3,314 USD / 342.7
Free Float	85.7%
Average Daily Value (mn)	127.96 USD
BofA Ticker / Exchange	LYFT / NAS
Bloomberg / Reuters	LYFT US / LYFT.OQ
ROE (2024E)	61.0%
Net Dbt to Eqty (Dec-2023A)	19.3%

BofA Securities does and seeks to do business with issuers covered in its research reports. As a result, investors should be aware that the firm may have a conflict of interest that could affect the objectivity of this report. Investors should consider this report as only a single factor in making their investment decision. Refer to important disclosures on page 7 to 9. Analyst Certification on page 5. Price Objective Basis/Risk on page 5.

12726867

Which one is the AI report?



Chipotle Mexican Grill: Sustained Growth and Innovation Drive BUY Rating

Reiterate Rating: BUY | PO: \$72.50 USD | Price: \$61.25 USD

Solid QX Performance Points to Strong QX 202X Trajectory

Chipotle's QX 202X performance delivered stronger-than-expected margins and solid same-store sales growth, supported by operational efficiencies and digital adoption. With a robust pipeline for QX 202X, including planned promotional activities and new product launches, we anticipate continued momentum.

QX 202X SSSG of 8.2% exceeded expectations of 7.5%, underpinned by transaction growth and a modest increase in average check size. Management's continued focus on labour optimization and cost controls has set a solid foundation for sustained profitability.

Chipotle Mexican Grill Big XQ traffic beat drives EPS upside; expect share gains to persist

Reiterate Rating: BUY | PO: 71.00 USD | Price: 51.78 USD

Strong XQ; expect XQ txns to recover post weather/tech

While CMG's XQ SSSG handily beat expectations on the topline (11% vs 8.4%/8.8% and the all-important transaction count (8.7% vs 4.5%/6.0%)), we believe investors are most focused on the deceleration through the quarter (XX was the strongest month) and softer start to XQ.

We continue to view the current demand environment as consistent with a slower – but still healthy – macroeconomy. As seasonality has shifted post-COVID, the summer in particular has been difficult to forecast (longer vacations/work from home) and we view the low end of the reiterated comp guide (mid to high single digits) as conservative.

Analyst report



Chipotle Mexican Grill Big 2Q traffic beat drives EPS upside; expect share gains to persist

Reiterate Rating: BUY | PO: 71.00 USD | Price: 51.78 USD

Strong 2Q; expect 3Q txns to recover post weather/tech

While CMG's 2Q SSSG handily beat expectations on the topline (11% vs 8.4%/8.8% BofAE/VA Consensus) and the all-important transaction count (8.7% vs 4.5%/6.0% BofAE/VA Consensus), we believe investors are most focused on the deceleration through the quarter (April was the strongest month) and softer start to 3Q. We continue to view the current demand environment as consistent with a slower – but still healthy – macroeconomy. As seasonality has shifted post-COVID, the summer in particular has been difficult to forecast (longer vacations/work from home) and we view the low end of the reiterated comp guide (mid to high single digits) as conservative. As mix grows less negative (more add-ons in 2Q vs 1Q), CMG brings back one of its most successful LTOs – smoked Brisket – in the fall, and advertising picks up we expect traffic to recover.

Labor guide looks conservative given throughput

Alongside robust SSSG trends, CMG posted record store margins – at 28.9% in 2Q (+140 bps y/y) — on sales leverage and more efficient ops. As CMG faces inflationary pressure from wages and food costs (protein, avocados), and invests behind improving portion consistency (a 40-60 bps headwind to COGS in 3Q), benefits from tightened operations (better labor deployment, dual-sided grill rollout) and ongoing throughput gains should help offset. Given leverage in 1H, the labor guide in particular could be high.

F24E EPS mostly unchg, 2Q beat offsets slightly lower 2H

Our F24E EPS moves to \$1.12 (vs \$1.13 prior) as BTE 2Q SSSG offsets lowered 3Q SSSG (5.9% vs 7.3% prior). Our estimate F24 SSSG of 7.5% (vs 7.1% previously) embeds 5.3% transaction growth and 2.5% average check growth.

Maintain PO and reiterate Buy rating

We value CMG on steady state earnings power, assuming 7000 U.S. stores and an additional 1000 international. At steady state, we expect AUVs to exceed \$4mm and margins to reach the prior peak of 27% (28% ex. pre-opening). Assuming G&A of 5% (similar to mature company operated systems), CMG would generate \$8.7bb in EBITDA. Applying a 24.5x multiple (unchanged), the implied EV is \$212 billion, or \$90bb discounted back to today, we derive our PO of \$71.

GPT-Model generated report



Chipotle Mexican Grill: Sustained Growth and Innovation Drive BUY Rating

Reiterate Rating: BUY | PO: \$68.32 USD | Price: \$61.25 USD

Solid Q4 Performance Points to Strong Q1 2025 Trajectory

Chipotle's Q4 2024 performance delivered stronger-than-expected margins and solid same-store sales growth (SSSG), supported by operational efficiencies and digital adoption. With a robust pipeline for Q1 2025, including planned promotional activities and new product launches, we anticipate continued momentum.

Q4 2024 SSSG of 8.2% exceeded expectations of 7.5%, underpinned by transaction growth and a modest increase in average check size. Management's continued focus on labor optimization and cost controls has set a solid foundation for sustained profitability.

Operational Efficiencies Driving Margins

Record restaurant-level margins in Q4 2024 (28.8%, +130bps YoY) position CMG for further upside despite ongoing inflationary pressures in protein and labor costs. Benefits from operational investments, such as the deployment of automated grills and streamlined digital order processing, continue to offset cost headwinds. For Q1 2025, margins are projected at 28.5%, with improvements in throughput efficiency and cost leverage.

2025 EPS Adjustments and Revenue Outlook

We have revised our 2025 EPS to \$1.28 (previously \$1.20) based on better-than-expected top-line performance and cost management. Revenues for 2025 are now forecasted at \$12.12 billion, with Q1 2025 expected to contribute approximately \$2.92 billion.

Maintain Price Objective and Buy Rating

Our valuation approach reflects steady-state earnings potential, assuming Chipotle reaches its long-term target of 8,000 stores globally. Applying a 25x EBITDA multiple to projected earnings and discounting back at a 7% WACC, we reaffirm our price objective of \$68.32, implying a ~12% upside from the current price.

Model Performance: Comparing Model vs Analyst Estimates



	Model forecast	Analyst forecast	% difference to analyst
Price target	68.32	71.00	-3.8%
2024E Revenue (m)	11,315	11,357	-0.4%
2025E Revenue (m)	12,120	12,976	-6.6%
2026E Revenue (m)	14,278	14,875	-4.0%
2024E Operating income (m)	1,955	2,062	-5.2%
2025E Operating income (m)	2,138	2,586	-17.3%
2026E Operating income (m)	2,582	3,065	-15.8%
2024E FCF (m)	1,571	1,464	7.3%
2025E FCF (m)	1,559	1,761	-11.5%
2026E FCF (m)	1,920	2,046	-6.2%
2024E EPS	1.15	1.12	2.7%
2025E EPS	1.28	1.36	-5.9%
2026E EPS	1.56	1.66	-6.0%

Limitations/Criticism

Results of little human guidance:

- Model lacks understanding for how a sell-side research analyst constructs a report
- Unrealistic assumptions when forecasting the DCF model
- GPT is inept at analysing complex excel spreadsheets.
- The first prompting stage is stable
- Model would generate different assumptions that heavily impacted results of DCF
- Requires human guidance

Unrealistic Stock price

Chipotle Mexican Grill Research Report

Investment Summary

**Reiterate Rating: BUY | Price Objective (PO):
\$2,020.00 | Current Price: \$1,780.00**



1. Introduction & Definition
2. Literature Review
3. Empirical Test on GPT For Equity Report
4. Model Evaluation
5. Conclusion

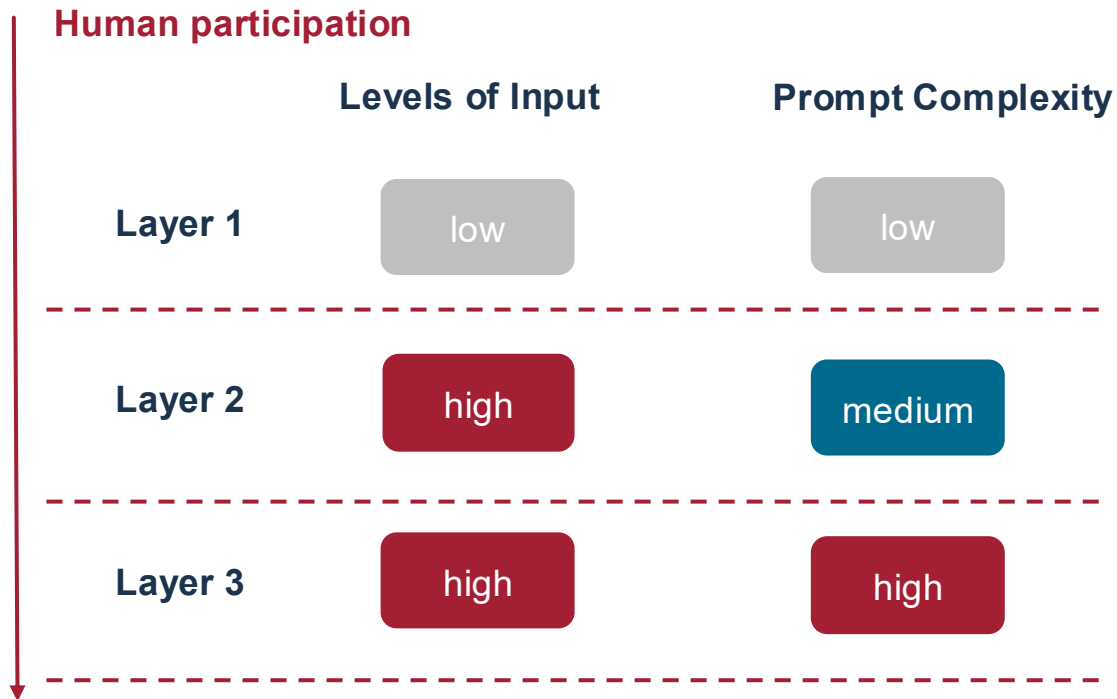
AGENDA



Evaluation Metrics



► Model Structure



► Evaluation metrics

- Each criterion focuses on a specific aspect of performance or perception

Criteria	Definition
Ease of Generation	How simple is the process of generating the report?
Stability	How stable are gen AI predictions (B/H/S)?
Forecasts comparison	How often does the AI rating (B/H/S) agree with analyst ratings?
Utility	Compared to the analyst report, which one strikes people as more useful?
Discernibility	Can people tell which one was AI generated?

Evaluation Metrics—Utility



- **Utility** is assessed through a breakdown of sub-criteria to ensure a thorough and objective evaluation of reports.

Table: Utility sub-criteria

Utility sub-criteria	Definition
Clarity and Logical Structure	Is the report well-organized, easy to understand, and follow a clear and logical flow?
Accuracy	Are the data, calculations, and conclusions correct and free from errors?
Comprehensiveness	Does the report cover all critical aspects?
Depth of Analysis	Does the analysis go beyond surface-level observations to provide deep insights with strong support and reasoning?
Actionability	Are the recommendations (B/H/S) clear, grounded in sound reasoning, and supported by robust model calculations with adequate explanation?

Table: Scoring Rubrics

Excellent (5)	Provides clear, actionable recommendations grounded in sound reasoning and robust models.
Good (4)	Recommendations are actionable and mostly well-supported, with minor gaps in reasoning.
Average (3)	Recommendations are present but lack sufficient clarity, justification, or explanation.
Poor (2)	Recommendations are impractical, or insufficiently explained.
Very Poor (1)	No actionable recommendations or entirely unsupported suggestions.

Evaluation Results



Ease of Generation

- **Layer 1: Simple**
~4 prompts, 80 words
- **Layer 2: Moderate**
~20 prompts, 700 words
- **Layer 3: Hard to train, easy to use**
~20 prompts (2000 words) for customization;
1 prompt for usage

Stability

Relatively stable with some variation.

Stability depends on prompts. Without clear constraints:
The model exhibits moderate variations in structure/sections, valuation methodologies, types of information analyzed, and final results.

Forecasts comparison

More granular prediction:
From Layer 1 to Layer 3, predictions become more granular.

- Layer1: Ratings
- Layer2: Adds return forecasts.
- Layer3: Incorporates price targets, calculated using DCF.

Discernibility

- AI integrates vast information efficiently, producing human-like reports.
- Limitations: lacks deep reasoning and understanding of complex interconnections.
- Differences are more pronounced in complex companies requiring subjective judgment.

Evaluation Results of Utility



- Analyst scores are higher than LLM models
- Progressive improvement in model scores (from Layer 1 to Layer 3)
- Chipotle (a pure-play company) shows greater score gains than MSFT (a more complicated company)

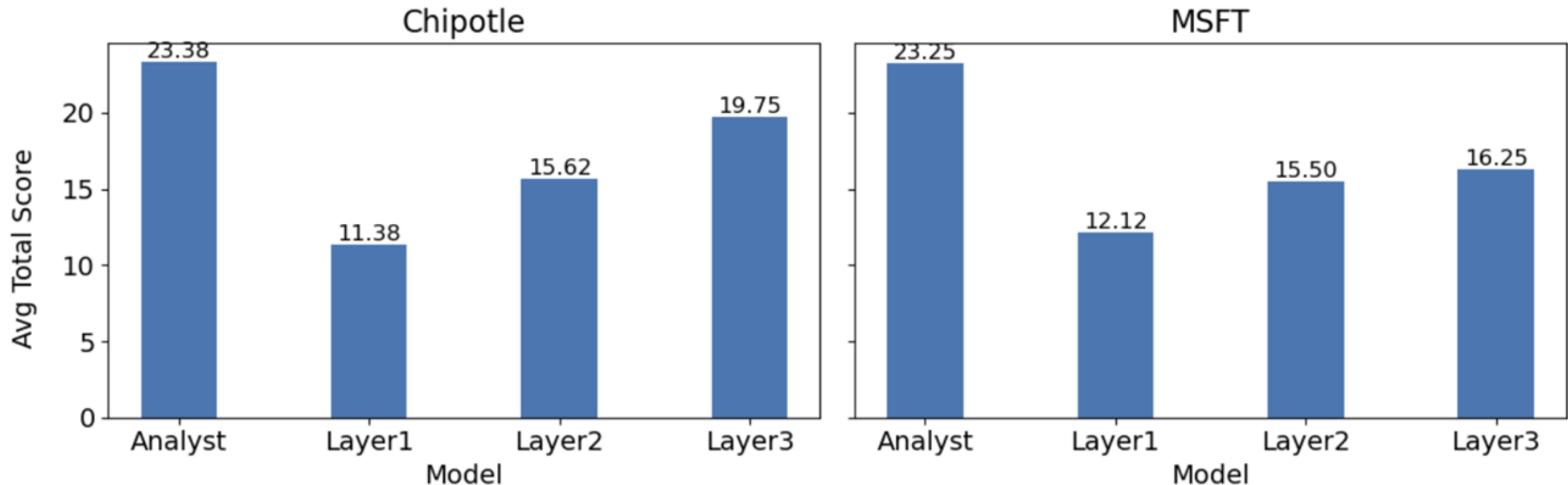
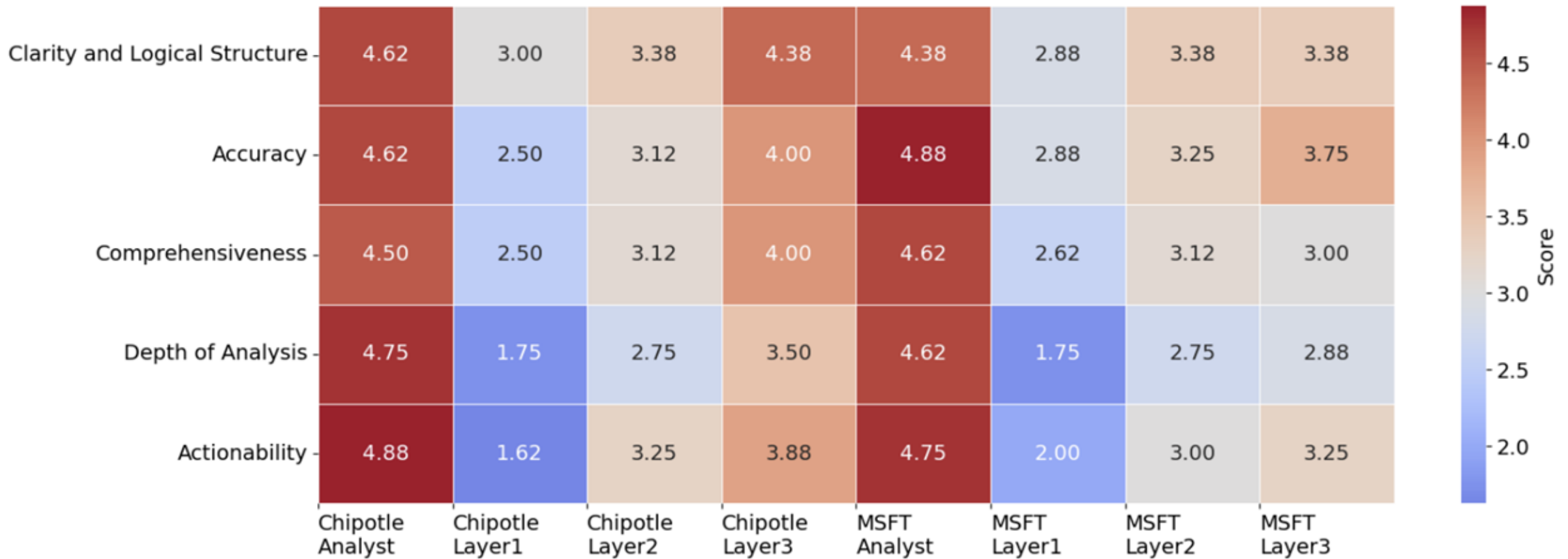


Table: Average Score of Reports (From 8 team members)

Evaluation Results of Utility



Total Scores Heatmap Across Models and Companies



Quantitative vs Qualitative data



Qualitative data

- Provides insights into industry trends, competitive landscapes, and regulatory environments. That allowed the model to be more forward looking.
- Including qualitative data significantly improves valuation accuracy. Leading to precise and reliable forecasts

Quantitative data

- Adjusts financial variables such as growth rates, discount rates, and margins (sensitivity analysis)
- Directly affects price targets and valuation results.
- Quantitative data alone results in less accurate DCF models.

Prompting

- Detailed prompting ensures precise numerical adjustments for sensitivity analysis, with inclusion of multiple valuation methods for robustness.
- Broad guidance may miss critical quantitative factors like growth rate variances.

Report examples



Quant Example

- **DCF Valuation:** *"Microsoft's historical revenue CAGR of 10% is used in a DCF model, with WACC at 8% and a terminal growth rate of 3%, resulting in a price target of \$470."*
- **Sensitivity Analysis:** *"Varying WACC between 7% and 9% showed a price range of \$440-\$500, demonstrating how quantitative inputs directly affect valuation."*

Qualitative example

- **Strategic Insights:** *"Microsoft's growth in AI through partnerships with OpenAI and strategic acquisitions (e.g., Activision Blizzard) creates a competitive advantage that underpins the bullish thesis."*
- **Risk Context:** *"Competitive pressures from AWS and Google Cloud limit Azure's potential, aligning with qualitative assessments of market dynamics."*

Natural Language Processing (NLP) Analysis



- We want to compare primary themes and lexical similarity between AI generated reports and an actual sell-side analyst report
- Leverage text **cosine similarity** to compare lexical similarity and Term Frequency-Inverse Document Frequency (**TF-IDF**) to perform a keyword thematic analysis

$$\text{TF-IDF}(t, d, D) = \text{TF}(t, d) \cdot \text{IDF}(t, D) \qquad \text{CosineSimilarity}(A, B) = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \cdot \sqrt{\sum_{i=1}^n B_i^2}}$$

NLP Analysis Results



Rank	Keyword (AI)	Weight (AI)	Keyword (Human)	Weight (Human)
1	metric	0.3668	eps	0.4777
2	price	0.3668	cmg	0.4777
3	cost	0.3437	usd	0.4246
4	usd	0.2445	sssg	0.2654
5	dec	0.2445	research	0.2654

66%

Degree of lexical overlap as
measured by cosine similarity

AI vs Human Sell-Side Research: Key Differences



Key Insights:

- **AI-Generated Report:**
 - Focus on generic terms: *metric, price, cost*.
 - Emphasis on quantitative, formulaic analysis.
 - Lacks industry-specific context and qualitative depth.
- **Human-Authored Report:**
 - Includes nuanced, industry-relevant terms: *sssg (same-store sales growth)*.
 - Blends quantitative data with qualitative insights: *research*.
 - Higher emphasis on critical financial metrics: *EPS*.
- **Cosine Similarity:**
 - Moderate overlap (**0.66**) due to common financial terms, but diverges in depth and specificity.

- 1. Introduction & Definition**
- 2. Literature Review**
- 3. Empirical Test on GPT For Equity Report**
- 4. Model Evaluation**
- 5. Conclusion**

AGENDA



Pros and Cons of GenAI



- Synthesizing large amounts of data
- Highly customizable to user requirements
- Strong performance in routine tasks

- **Inconsistent with complex business models**
- **Inaccuracy**
- **Lack of reasoning and granular understanding of the underlying business**

An aerial photograph of a city skyline across a wide river. In the foreground, a multi-lane road curves along a green park area with trees showing autumn foliage. A bridge is visible in the middle ground, crossing the river. The city skyline in the background features numerous skyscrapers under a clear blue sky with scattered clouds. The text "Thank you! Q&A" is overlaid in large white font on the right side of the image.

Thank you!
Q&A