FinRobot: AI Agent for Equity Research and Valuation with Large Language Models

Tianyu Zhou   Pinqiao Wang   Yilin Wu   Hongyang Yang[†]

AI4Finance Foundation

11/14/2024

AI in Finance

GenAI is becoming more capable of reasoning
People need automation support in Financial Analysis, but GenAI (LLM) produces results that are... too artificial.

> Why did Apple's Q3 net income drop despite higher sales this year?

Apple's Q3 net income drop could be due to multiple factors, even with increased sales. Some of these factors might include:

1. Increased expenses: Apple might have incurred higher operational costs, like R&D or marketing.

2. Currency fluctuations: Exchange rates could have negatively impacted profits, particularly if the dollar is strong.

3. Product mix: If sales of lower-margin products increased, overall profitability could suffer.

4. Supply chain issues: Increased costs due to supply chain disruptions could reduce net income.

These are potential areas that might explain the net income drop, though for precise reasons, one would need to look at Apple's detailed financial reports and any statements from the company.

Related Works in Financial AI

Large Language Models (LLMs) in Finance

- Applications:
    - Sentiment analysis for gauging market sentiment.
    - Market prediction for forecasting trends.
- Limitations:
    - Lack of real-time adaptability, limiting responsiveness.
    - Limited sector-specific expertise, reducing accuracy in equity research.

AI Agents with Multi-Agent Systems

- Examples: FinAgent, FinMem
- Capabilities:
    - Use real-time data to inform trading strategies.
    - Chain-of-Thought (CoT) prompting allows step-by-step reasoning for predictions.
- Limitations:
    - Limited to trading applications, lacking depth for comprehensive equity research.

Introduction

FinRobot:
First AI Agent for Equity Research Using a Multi-Agent Chain of Thought (CoT) System

- Delivers actionable insights at professional-level of a human analyst
- Blending discretionary judgment with real-time data and new evaluation metrics
- And it is Open-Sourced
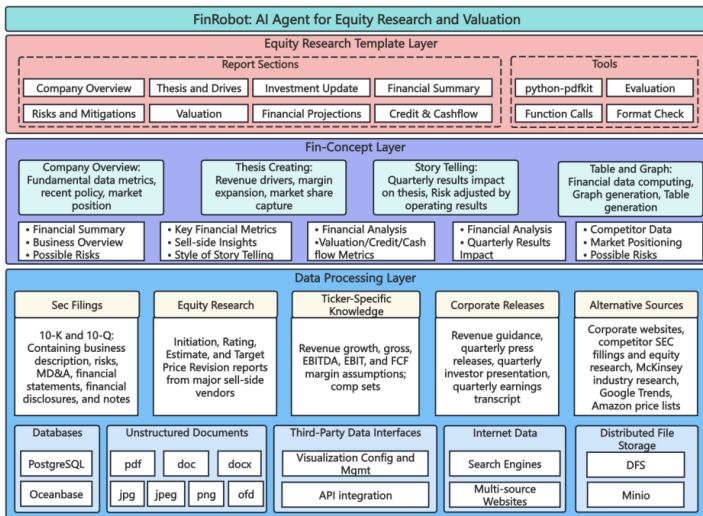

FinRobot

Main Framework

Figure 1: Overall Framework of FinRobot.

Multi-Agent System

FinRobot CoT
Multi-Agents Interaction throughout different layers

- Data-CoT Agent:
  - Aggregates sector-specific data for a robust financial foundation.
- Concept-CoT Agent:
  - Emulates human analyst reasoning to create actionable insights.
- Thesis-CoT Agent:
  - Synthesizes insights into a structured, coherent investment thesis and report.

Key Advantages of FinRobot
- Provides real-time adaptability, allowing it to respond to market changes dynamically.
- Delivers nuanced, in-depth analysis akin to professional brokerage firms.
- Addresses traditional AI limitations, increasing both accuracy and practical relevance in equity research.
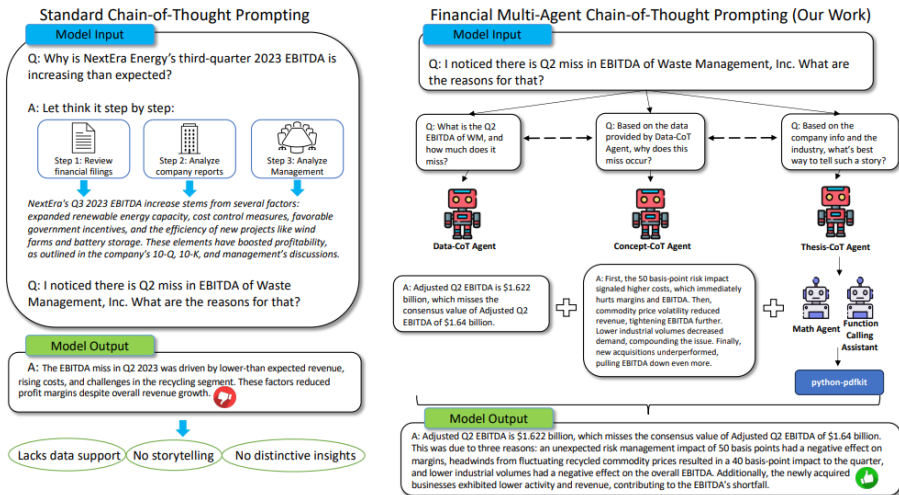
Financial Multi-Agent CoT comparison

Figure 2: Standard Chain-of-Thought (CoT) vs Financial Multi-Agent Chain-of-Thought (CoT)

Latest Results

US EQUITY RESEARCH

7 Oct 2024

WASTE MANAGEMENT, INC. (WM)

Waste Management (WM) demonstrates steady growth with a 7.2% EBITDA CAGR, a contracting 1.7% dividend yield, and faces valuation challenges, risks from fluctuating commodity prices, regulatory changes, and high leverage from acquisitions, all while maintaining fair value per share projections between $144.3 and $176.6 based on improved EBITDA margins and pricing power in recent years.

Company Overview

Waste Management (WM), the largest and most prominent player in its sector, presents an intriguing investment proposition. An analysis of the CAGR of its EBITDA reveals a growth rate of 7.2%, suggesting that though the company is exhibiting solid performance, the growth is somewhat limited. Multiple factors could account for this limitation, including sector-specific challenges or strategic decisions made by the company. Furthermore, the company's dividend yield has contracted to 1.7%, a modest figure attributed to the stock price appreciating in recent years. This could indicate improving investor confidence, but also suggests lower relative returns for dividend investors.

Investment Update

Adjusted Q2 EBITDA is $1.622 billion, which misses the consensus value of Adjusted Q2 EBITDA of $1.64 billion. This was due to three reasons: an unexpected risk management impact of 50 basis points had a negative effect on margins, headwinds from fluctuating recycled commodity prices resulted in a 40 basis-point impact to the quarter, and lower industrial volumes had a negative effect on the overall EBITDA. Additionally, the newly acquired businesses exhibited lower activity and revenue, contributing to the EBITDA's shortfall.

Valuation

The fair value per share is projected to range between $144.3 and $176.6, essentially based on the EV/EBITDA ranges of 13x – 15x from 2019 to 2023. The average EV/EBITDA from 2019 to 2023 is observed to be 14x, which significantly exceeds the historical average of 6x from 2004 to 2015, attributable to the improved pricing power in the industry between 2015 and 2017. Furthermore, during the course of 2019 to 2023, the revenue demonstrated an average increase of 6.7%, while the EBITDA margin consistently climbed, falling slightly in 2020 but eventually escalating to 28.9% by 2023.

Risks

Risks include (1) fluctuations in commodity prices potentially affecting revenue, (2)

Key Financial Data

Bloomberg Ticker	WM US
Sector	Utilities (Environmental Services)
Share Price (USD)	207.32
Rating	BUY
12-mth Target Price (USD)	219.17
Market Cap (USDb)	86.48
Volume (m shares)	1.43
Free float (%)	99.10
Dividend yield (%)	1.62
Net Debt to Equity (%)	172.50
Fwd. P/E (x)	27.5
P/Book (x)	8.5
ROE (%)	29.10

Evaluation

We engaged a panel of investment banking analysts to review and rate the generated report.

Score	Accuracy	Logicality	Storytelling
10	Perfect accuracy, no errors or inconsistencies.	Exceptionally logical flow, perfectly structured.	Exceptionally engaging, flawless storytelling.
9	Highly accurate with only minor errors.	Very logical, well-structured with minor flaws.	Strong storytelling, minor areas for improvement.
8	Mostly accurate, a few minor inconsistencies.	Logical and mostly well-structured, minor issues.	Engaging and well-paced, slight room for improvement.
7	Accurate, with a few minor issues.	Generally logical with clear structure.	Good storytelling, clear and engaging.
6	Satisfactory accuracy, some minor errors present.	Mostly logical but may lack depth in reasoning.	Clear storytelling, but lacks depth or consistency.
5	Fairly accurate but noticeable issues.	Basic logical structure, some gaps in flow.	Basic storytelling, lacks engagement.
4	Somewhat accurate but may contain errors.	Lacks consistent logical structure, some confusion.	Unengaging storytelling, basic and inconsistent.
3	Contains frequent inaccuracies.	Poor logical flow, structure is disjointed or unclear.	Storytelling is unclear, lacks engagement.
2	Inadequate accuracy, many errors.	Very little logical structure, mostly confusing.	Poor storytelling, lacks clarity and coherence.
1	Highly inaccurate, numerous errors.	Little to no logical flow, very confusing structure.	Extremely poor storytelling, difficult to follow.
0	Completely inaccurate, no correct information.	No logical structure, completely incoherent.	No storytelling structure, entirely confusing.

Table 1: Grading Criteria

Evaluation

Evaluation Results:

- FinRobot received high scores from investment banking analysts across three dimensions:
  - Accuracy: 9.5
  - Logical Coherence: 9.2
  - Storytelling: 8.3

Key Insights:

- Strong factual precision noted; slight variation in logical coherence scores.

- Storytelling received positive feedback, with suggestions to enhance narrative engagement.

Reviewer	Accuracy	Logicality	Storytelling
Reviewer 1	10	10	10
Reviewer 2	10	9	8
Reviewer 3	10	9	8
Reviewer 4	9	9	7
Reviewer 5	9	10	7
Reviewer 6	9	9	10
Reviewer 7	10	9.5	8.5

Table 2: Scores for Equity Research Report by Seven Reviewers

Evaluation

Compared with other prompting techniques

- Generated multiple reports on the same topic.
- Evaluated by simulation for consistency in:
    - Accuracy
    - Logicality
    - Storytelling
- Compared with zero-shot, few-shot, and chain-of-thought prompting methods.
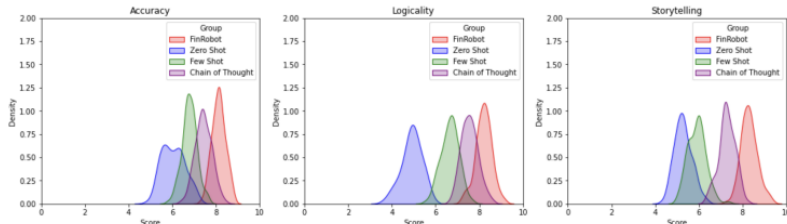

Figure 4: Quality Analysis

Future Work

Broader Scope:

- Expand coverage to all 11 sectors.
- Include comprehensive reports for NASDAQ-100 companies.
- Collaborate with major financial institutions to generate equity research reports.

Transparent FinRobot:

- Explainable, transparent FinRobot for enhanced trust and usability.

Jeffrey S. Abarbanell and Brian J. Bushee. 1997. Fundamental analysis, future earnings, and stock prices. *Journal of Accounting Research*, 35(1), 1–24.

Karen Berman and Joe Knight. 2013. *Financial Intelligence, Revised Edition: A Manager's Guide to Knowing What the Numbers Really Mean*. Harvard Business Review Press.

Tom Brown, Benjamin Mann, Nick Ryder, et al. 2020. Language models are few-shot learners. *Advances in Neural Information Processing Systems*, 33, 1877–1901.

B.C. Greenwald, J. Kahn, P.D. Sonkin, and M. van Biema. 2004. *Value Investing: From Graham to Buffett and Beyond*. Wiley.

Bruno Miranda Henrique, Vinicius Amorim Sobreiro, and Herbert Kimura. 2019. Literature review: Machine learning techniques applied to financial market prediction. *Expert Systems with Applications*, 124, 226–251.

Allen H. Huang, Hui Wang, and Yi Yang. 2023. FinBERT: A large language model for extracting information from financial text. *Contemporary Accounting Research*, 40(2), 806–841.

Weiwei Jiang. 2021. Applications of deep learning in stock market prediction: recent progress. *Expert Systems with Applications*, 184, 115537.

Alex Kim, Maximilian Muhn, and Valeri V. Nikolaev. 2024. Financial Statement Analysis with Large Language Models. *Chicago Booth Research Paper* Forthcoming, Fama-Miller Working Paper.

Pinqiao Wang
Columbia University
pw2594@columbia.edu

Questions?