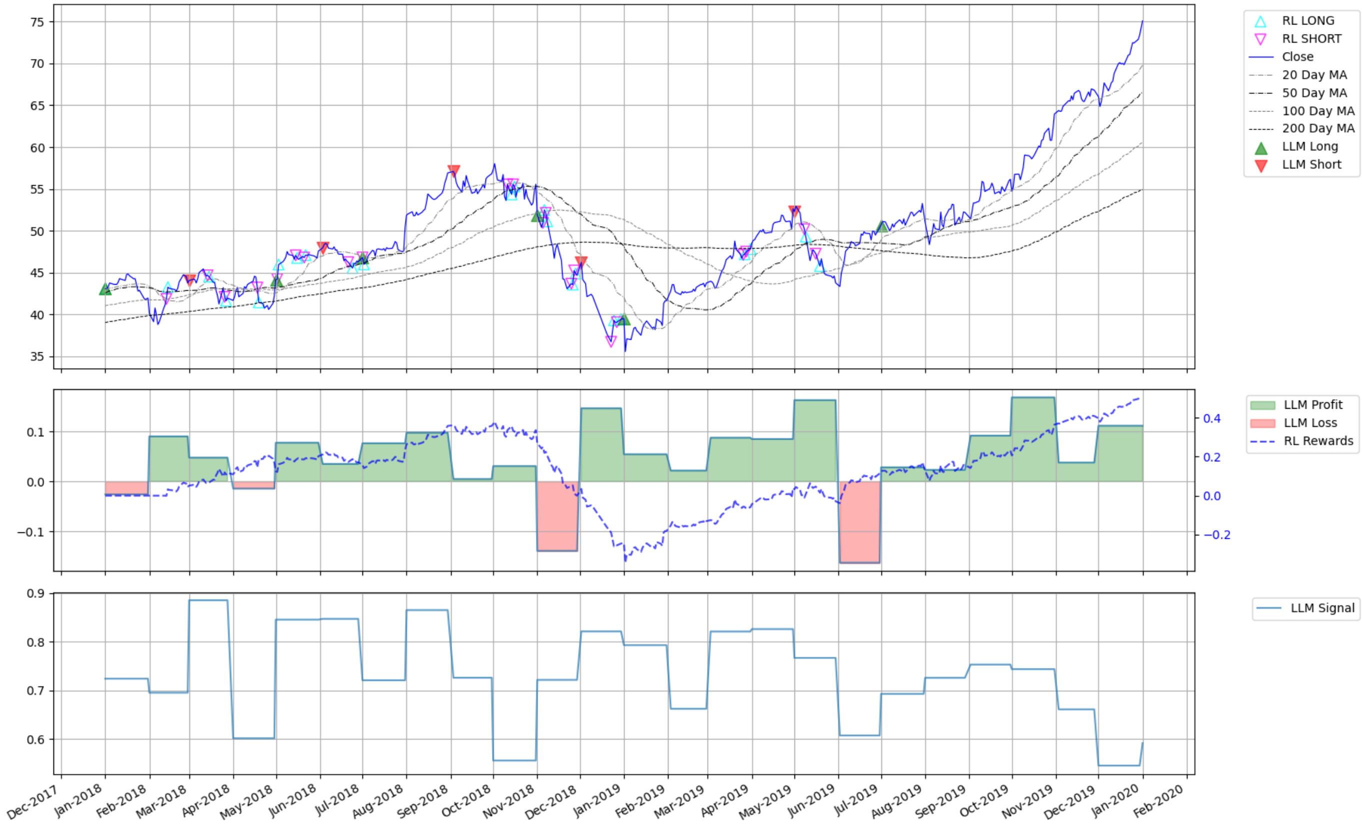

Fig. 2. AAPL Performance with LLM+RL Model.

TABLE IX
EXPERIMENT 2 RESULTS: SHARPE RATIO

Ticker	LLM+RL (σ)	RL-Only (σ)	LLM-Only
AAPL	1.70 (0.43)	1.42 (0.05)	**2.09**
AMZN	**1.21 (0.58)**	0.42 (0.23)	0.84
GOOGL	**1.16 (0.17)**	0.23 (0.37)	1.12
META	0.46 (0.75)	0.15 (0.61)	**0.77**
MSFT	**1.16 (0.28)**	0.99 (0.30)	0.50
TSLA	**0.92 (0.19)**	0.62 (0.60)	0.87
Mean	**1.10**	0.64	1.03

TABLE X
EXPERIMENT 2 RESULTS: MAXIMUM DRAWDOWN

Ticker	LLM+RL (σ)	RL-Only (σ)	LLM-Only
AAPL	0.29 (0.20)	0.45 (0.01)	**0.28**
AMZN	0.26 (0.12)	**0.19 (0.14)**	0.34
GOOGL	0.28 (0.06)	**0.25 (0.18)**	0.35
META	0.35 (0.11)	0.45 (0.27)	**0.30**
MSFT	0.19 (0.08)	**0.17 (0.09)**	0.21
TSLA	**0.46 (0.05)**	0.65 (0.13)	0.59
Mean	**0.31**	0.36	0.35

The hybrid agent did not consistently minimize MDD per stock but achieved values close to the best across agents, with the lowest overall mean (0.31). This suggests overall smoother drawdowns under uncertainty across the universe

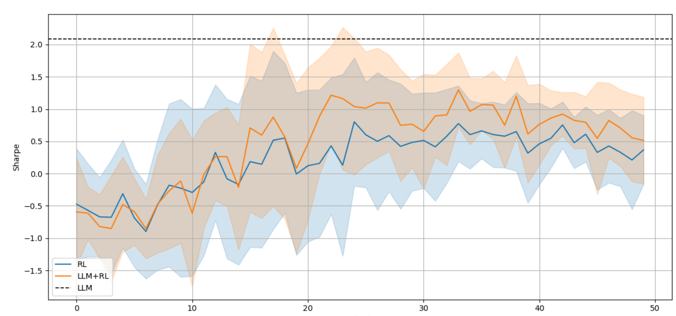

Fig. 3. Training Behavior for AAPL: Sharpe Ratio.

(see Tables IX and X).

IV. CONCLUSION AND FUTURE WORK

This study has explored an RL+LLM hybrid architecture for algorithmic trading, where LLMs generate guidance for RL agents to act as tactical executors.

Experiment 1 has shown that well engineered prompts improve the LLM's performance, with Prompt 4 achieving the highest SR and lowest uncertainty. Expert evaluations confirmed the rationale of generated strategies within the domain.

Experiment 2 has demonstrated that an RL agent guided by LLM signals outperforms the RL-only baseline in four out of

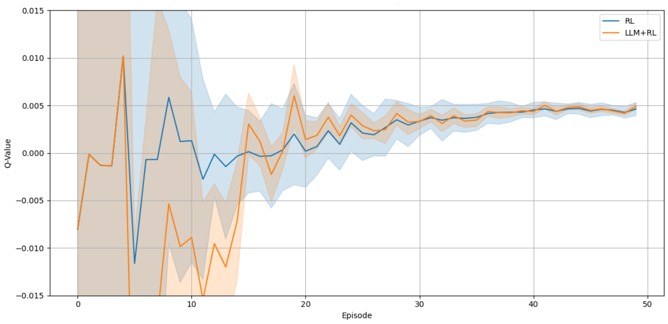

Fig. 4. Training Behavior for AAPL: Q-Values for LONG.

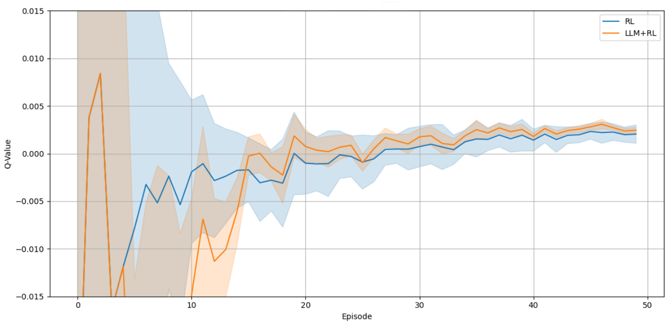

Fig. 5. Training Behavior for AAPL: Q-Values for SHORT.

six stocks when evaluated by their Sharpe Ratio. While MDD was not consistently reduced, the overall drawdowns remained low on average. Importantly, the underlying RL architecture was not modified; all observed improvements stemmed from LLM guidance.

Future research should address two main directions. First, while the LLM can guide the RL, reward shaping is necessary to attain optimal results. Second, modular specialization through multiple LLM agents prompted for specific domains may enable a mixture-of-experts architecture, and lessen the risk of confabulation.

Overall, this work presents a novel LLM+RL system that improves both return and risk outcomes. It supports modular, agentic setups where LLMs operate as trustworthy planners in financial decision making.

Supplementary Material

Full prompt templates (strategy and analyst), labeling-heuristic pseudocode, extended dataset schema, and complete replication tables are available from the corresponding author upon request.

Acknowledgment

We thank the expert reviewers who contributed their time and expertise to this work.

References

[1] M. Chlistalla, "High-frequency trading: Better than its reputation?" Deutsche Bank Research, Frankfurt am Main, Germany, Tech. Rep. Research Briefing, Feb. 2011. [Online]. Available: https://www.finextra.com/finextra-downloads/featuredocs/prod0000000000269468.pdf

[2] E. F. Fama and K. R. French, "A five-factor asset pricing model," *Journal of Financial Economics*, vol. 116, no. 1, pp. 1–22, 2015.

[3] T. Théate and D. Ernst, "An application of deep reinforcement learning to algorithmic trading," *Expert Systems with Applications*, vol. 173, p. 114632, Jul. 2021.

[4] S. M. Bartram, J. Branke, and M. Motahari, *Artificial Intelligence in Asset Management*. CFA Institute Research Foundation, 2020.

[5] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, 2nd ed., ser. Adaptive Computation and Machine Learning series. MIT Press, 2018. [Online]. Available: http://incompleteideas.net/book/the-book-2nd.html

[6] X.-Y. Liu, H. Yang, J. Gao, and C. D. Wang, "Finrl: Deep reinforcement learning framework to automate trading in quantitative finance," in *Proceedings of the Second ACM International Conference on AI in Finance*. ACM, Nov. 2021.

[7] M. M. L. de Prado, "Beyond econometrics: A roadmap towards financial machine learning," *Econometric Modeling: Theoretical Issues in Microeconometrics eJournal*, 2019. [Online]. Available: https://api.semanticscholar.org/CorpusID:199365784

[8] L. Onozo, F. Arthur, and B. Gyires-Tóth, "Leveraging LLMs for financial news analysis and macroeconomic indicator nowcasting," *IEEE Access*, 2024, early Access. Online. Accessed: Feb. 10, 2025.

[9] W. Zhang *et al.*, "A multimodal foundation agent for financial trading: Tool-augmented, diversified, and generalist," New York, NY, USA, p. 4314–4325, 2024.

[10] M. Wang, K. Izumi, and H. Sakaji, "LLMFactor: Extracting profitable factors through prompts for explainable stock movement prediction," 2024. [Online]. Available: https://arxiv.org/abs/2406.10811

[11] S. Schulhoff *et al.*, "The prompt report: A systematic survey of prompting techniques," 2024. [Online]. Available: https://arxiv.org/abs/2406.06608

[12] Y. Yu *et al.*, "Finmem: A performance-enhanced llm trading agent with layered memory and character design," in *Proceedings of the AAAI Spring Symposium Series*, R. P. A. Petrick and C. W. Geib, Eds. AAAI Press, Jan. 2024, pp. 595–597.

[13] K. Arulkumaran, M. P. Deisenroth, M. Brundage, and A. A. Bharath, "Deep reinforcement learning: A brief survey," *IEEE Signal Processing Magazine*, vol. 34, no. 6, pp. 26–38, 2017.

[14] A. Lopez-Lira and Y. Tang, "Can chatgpt forecast stock price movements? return predictability and large language models," 2023.

[15] J. Huang and K. C.-C. Chang, "Towards reasoning in large language models: A survey," 2022. [Online]. Available: https://arxiv.org/abs/2212.10403

[16] M. Lopez de Prado, "The 10 reasons most machine learning funds fail," The Journal of Portfolio Management, Tech. Rep., 06 2018.

[17] L. Takara, A. Santos, V. Mariani, and L. Coelho, "Deep reinforcement learning applied to a sparse-reward trading environment with intraday data," *Expert Systems with Applications*, vol. 238, p. 121897, 2024.

[18] A. Chaddha and S. Yadav, "Examining the predictive power of moving averages in the stock market," *Journal of Student Research*, vol. 11, no. 3, 2022.

[19] S. Wang, H. Yuan, L. M. Ni, and J. Guo, "Quantagent: Seeking holy grail in trading by self-improving large language model," 2024. [Online]. Available: https://arxiv.org/abs/2402.03755

[20] A. Lopez-Lira, Y. Tang, and M. Zhu, "The memorization problem: Can we trust LLMs' economic forecasts?" 2025. [Online]. Available: https://arxiv.org/abs/2504.14765

[21] H. Gonen, S. Iyer, T. Blevins, N. Smith, and L. Zettlemoyer, "Demystifying prompts in language models via perplexity estimation," in *Findings of the Association for Computational Linguistics: EMNLP 2023*, H. Bouamor, J. Pino, and K. Bali, Eds. Singapore: Association for Computational Linguistics, Dec. 2023, pp. 10 136–10 148.

[22] A. Kaltchenko, "Entropy heat-mapping: Localizing GPT-based OCR errors with sliding-window shannon analysis," 2025. [Online]. Available: https://arxiv.org/abs/2505.00746

[23] L. M. Demajo, V. Vella, and A. Dingli, "Explainable AI for interpretable credit scoring," 2020. [Online]. Available: https://arxiv.org/abs/2012.03749

[24] G. Yona, R. Aharoni, and M. Geva, "Can large language models faithfully express their intrinsic uncertainty in words?" 2024. [Online]. Available: https://arxiv.org/abs/2405.16908

Appendix A
Strategy Prompt

The final tuned prompt from Experiment 1 and the LLM strategy generator for Experiment 2, is available in 1.

Listing 1. Tuned Strategy Prompt

```
User_Context:
  Last_Strategy_Used_Data:
    last_returns: "{Last_LLM_Strat_Returns}"
    last_action: "{Last_LLM_Strat_Action}"
    Rationale: |
       """{Last_LLM_Strat}"""


  Stock_Data:
    General:
      Beta: {Market_Beta}
      Classification: {classification}

    Last_Weeks_Price:
      Close: "{Close}"
      Volume: "{Volume}"

    Weekly_Past_Returns: "{Weekly_Past_Returns}"

    Historical_Volatility:
      HV_Close: "{HV_Close}"

    Implied_Volatility:
      IV_Close: "{IV_Close}"

  Fundamental_Data:
    Ratios:
      Current_Ratio: "{Current_Ratio}"
      Quick_Ratio: "{Quick_Ratio}"
      Debt_to_Equity_Ratio: "{Debt_to_Equity_Ratio}"
      PE_Ratio: "{PE_Ratio}"
    Margins:
      Gross_Margin: "{Gross_Margin}"
      Operating_Margin: "{Operating_Margin}"
      Net_Profit_Margin: "{Net_Profit_Margin}"
    Growth Metrics:
      EPS_YoY: "{EPS_YoY_Growth}"
      Net_Income_YoY: "{Net_Income_YoY_Growth}"
      Free_Cash_Flow_YoY: "{Free_Cash_Flow_Per_Share_YoY_Growth}"

  Technical_Analysis:
    Moving_Averages:
        20MA: "{20MA}"
        50MA: "{50MA}"
        200MA: "{200MA}"
    MA_Slopes:
        20MA_Slope: "{20MA_Slope}"
        50MA_Slope: "{50MA_Slope}"
        100MA_Slope: "{100MA_Slope}"
        200MA_Slope: "{200MA_Slope}"
    MACD:
        Value: "{MACD}"
        Signal_Line: "{Signal_Line}"
        MACD_Strength: {MACD_Strength}
    RSI:
        Value: "{RSI}"
    ATR: "{ATR}"

  Macro_Data:
    Macro_Indices:
      SPX:
        Close: "{SPX_Close}"
        Close_20MA: "{SPX_Close_MA}"
        Close_Slope: "{SPX_Close_Slope}"
      VIX:
        Close: "{VIX_Close}"
        Close_20MA: "{VIX_Close_MA}"
        Close_Slope: "{VIX_Close_Slope}"
    Economic_Data:
      GDP_QoQ: "{GDP_QoQ}"
      PMI: "{PMI}"
      Consumer_Confidence_QoQ: "{Consumer_Confidence_QoQ}"
      M2_Money_Supply_QoQ: "{M2_Money_Supply_QoQ}"
      PPI_YoY: "{PPI_YoY}"
      Treasury_Yields_YoY: "{Treasury_Yields_YoY}"

  Options_Data:
    Put_IV_Skews:
      OTM_Skew: "{OTM_Skew}"
      ATM_Skew: "{ATM_Skew}"
      ITM_Skew: "{ITM_Skew}"
    20Day_Moving_Averages:
      OTM_Skew_MA: "{MA_OTM_Skew}"
      ATM_Skew_MA: "{MA_ATM_Skew}"
      ITM_Skew_MA: "{MA_ITM_Skew}"

  News_Sentiment: {news_sentiment}
  News_Impact_Score: {news_impact_score}

System_Context(System):
  Persona: {persona}
  Portfolio_Objectives: {portfolio_objectives}
  Instructions: |
    Develop a LONG or SHORT trading strategy for a single stock only for the next Month that aligns with the `portfolio_objectives`. Follow these guidelines:

    1. Stock Analysis:
       - Evaluate price trends: Compare the Close price against 20MA, 50MA, and 200MA to assess momentum or reversals.
       - Analyze returns: Use Weekly Past Returns to validate trend sustainability.
       - Contextualize volatility: Align `HV_Close` and `HV_High` with recent price action for trend validation.
       - Incorporate beta: Use `beta` to gauge sensitivity to market movements.

       - ICL Example: "Close␣price␣above␣20MA␣and␣50MA␣with␣steep␣20MA␣slope␣signals␣bullish␣momentum.␣Weekly␣returns␣confirm␣a␣sustainable␣uptrend."

    2. Technical Analysis:
       - Use RSI: Identify momentum signals (>70 overbought; <30 oversold) and divergences for reversals.
       - Validate with `MACD`: Use crossovers of `MACD.Value` and `Signal_Line`, and `MACD_Strength` for directional confidence.
       - Leverage `RSI.value` divergences, and steep `Moving_Averages` slopes. Or focus on stable `Moving_Averages` patterns on stable historical volatility `HV_Close`.
       - ICL Example: "RSI␣at␣65,␣a␣positive␣MACD␣crossover␣indicate␣bullish␣momentum."

    3. Fundamental Analysis:
       - Evaluate growth metrics: Use `EPS_YoY`, `Net_Income_YoY`, and `Free_Cash_Flow_YoY` for profitability and sustainability.
       - Prioritize ratios: Low `Debt_to_Equity_Ratio` and `Current_Ratio` reflect financial stability.
       - Focus on aggressive `Growth Metrics` and earnings news.
       - ICL Example: "EPS_YoY␣growth␣of␣25%␣and␣low␣Debt-to-Equity␣ratio␣of␣0.5␣support␣strong␣financial␣health,␣aligning␣with␣a␣LONG␣strategy."

    4. Macro Analysis:
      - Align with market sentiment across `Macro_Data`:
        - "SPX_Close_Slope␣>␣0␣&&␣VIX_Close_Slope␣<␣0": Bullish (Risk-On)
        - "SPX_Close_Slope␣<␣0␣&&␣VIX_Close_Slope␣>␣0": Bearish (Risk-Off)
      - Validate with `Economic_Data`:
        - "GDP_QoQ␣>␣0␣&&␣`PMI`␣>␣50" leads to Economic Expansion
        - "`Treasury_Yields_YoY`␣<␣0" Signals Recession Risk, especially if already
```