Risko1: Reasoning for Risk Management Governed by Structural Constraints

Marco Molinari*
LSE.AI
London School of Economics
m.molinari1@lse.ac.uk

Deasy Darlene Tunas*
LSE.AI
London School of Economics
d.d.tunas@lse.ac.uk

Saharsha Navani*
Department of Computer Science
Princeton University
snavani@princeton.edu

Elliot Gao
LSE.AI
London School of Economics
t.gao10@lse.ac.uk

Abstract

Risk management is a crucial task for both individual investors and financial institutions that seek to identify and quantify the risks they are exposed to. We introduce *Risko1*, an 8B-parameter financial reasoning model trained with Group Relative Policy Optimization (GRPO) on both textual context and financial information. It identifies specific risks to which companies are exposed and quantifies their impact in terms of standard risk metrics: Value at Risk (VaR), Conditional Value at Risk (CVaR), and Volatility. These metrics must strictly satisfy fundamental constraints such as $\mathrm{CVaR}\alpha > \mathrm{VaR}\alpha$ and monotonicity across confidence levels. Performance is slightly above the much larger Llama 3.3 70B in accuracy, and roughly on par in Mean Squared Error (MSE). Beyond quantitative ability, we analyze the quality of the risk scenarios generated. Regulators require institutions to establish controls to mitigate risk exposure. This is done using a risk taxonomy that classifies risks across tiers based on granularity, with 1 being a broad category (for instance, operational risks), and 4 being the most granular (a specific event). Controls are enacted at the appropriate level of granularity. We explore the distribution of tiers of generated risks, and find that they are coherent with the given context (mainly market and operational) and granular (mostly tier 3 and 4), and hence amenable to mitigation, as controls may be assigned effectively. [2]

1 Introduction

Risk numbers determine returns, capital requirements, and solvency: VaR, CVaR, and volatility. When those numbers diverge or contradict fundamental constraints, desks overreact, capital is misallocated, and trust erodes [39, 40] with potentially disastrous consequences [24]. For stress testing, institutions need families of scenarios with different levels of severity, which AI generates. Investors and portfolio managers face the same tension at a different scale: they seek exposure to known, compensated risks to realize returns, but want to avoid unexpected losses as much as possible.

We focus on companies in the Dow Jones Industrial Average (DJIA): given financials and company context from the FinGPT dataset [37], we identify concrete risk scenarios and produce VaR, CVaR,

*Equal contributions co-first author.

[2]Code for **all** experiments, data, and weights available at: `https://github.com/FlexCode29/financial_grpo_reproducibility_code`.

39th Conference on Neural Information Processing Systems (NeurIPS 2025) Workshop: Generative AI in Finance.

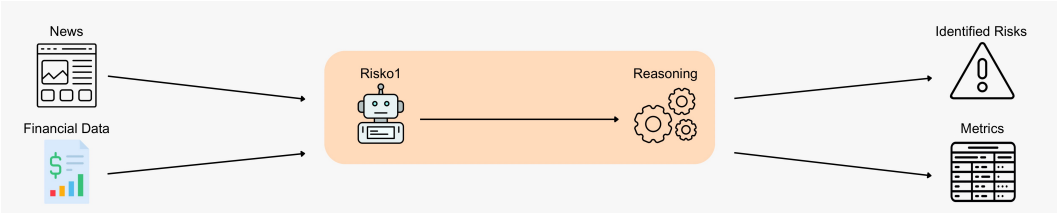

Figure 1: System level overview of Risko1.

and volatility that obey constraints like $\mathrm{CVaR} > \mathrm{VaR}$ and monotonicity across confidence levels. This prevents distortion as downstream systems may consume the numbers directly.

Large Language Models (LLMs) [35] have seen adoption across sectors, including finance [20, 28]. However, we find that specialized financial LLMs [28] can break constraints, while general purpose LLMs are compute hungry, add latency to reactive systems, and can fall behind on key metrics.

We introduce **Risko1**, an 8B financial reasoner trained with LoRA and GRPO to optimize for accuracy and constraint satisfaction. The model reads financial context, generates actionable risk scenarios that map into a risk taxonomy, and returns VaR, CVaR, and volatility accurately. On held out data, Risko1 matches Llama 3.3 70B on MSE and slightly exceeds it in accuracy while maintaining near perfect constraint compliance. Our contributions can be summarized as follows:

- We develop the Risko1 reasoning model for risk scenario generation (Figure 1).
- We evaluate scenarios generated by Risko1 on accuracy and MSE with results on par or above Llama 3.3 70b. They also prove coherent with the context and suitable for mitigation.

2 Background

2.1 GenAI in finance

Generative AI (GenAI) has found applications in the financial domain [29, 25, 11, 17]. However, generalist LLMs often struggle with the unique demands of financial reasoning [7, 17], which requires domain knowledge [28, 38]. Hence, recent research has focused on developing specialized models trained on domain data [9]. Notable contributions in this area include the **Fin-R1** model [20], and later the improved **Fin-o1** model [28], which is a fine-tune of Llama 3.1 8B [13]. Both works construct high-quality, Chain-of-Thought (CoT) datasets (FinCoT and Fin-R1-Data, respectively). Later, they employ a two-stage training process, beginning with Supervised Fine-Tuning (SFT) for domain adaptation, followed by Reinforcement Learning (RL) to refine the model's reasoning [19, 30, 8].

2.2 Reasoning with GRPO

The development of models like Fin-o1 and Fin-R1 was made possible by a prior shift towards foundational reasoning models, for instance [10], which demonstrated improved capabilities on complex tasks [22]. Reinforcement Learning (RL) has proved a significant improvement over simple Supervised Fine-Tuning (SFT). In particular, group Relative Policy Optimization (GRPO) [31], is an efficient RL algorithm that eliminates the need for a critic by reformulating the advantage for a prompt as a response's reward relative to its peers within a group of G responses. GRPO's objective function consists of this advantage and a Kullback–Leibler (KL) divergence penalty for stability:

$$\mathcal{J}_{\text{GRPO}}(\theta) = \hat{\mathbb{E}}_{q\sim\mathcal{D},\{o_i\}\sim\pi_{\theta_{\text{old}}}}\left[\frac{1}{G}\sum_{i=1}^{G} L^{\text{CLIP}}(\theta, o_i, \hat{A}_i)\right] - \beta D_{KL}(\pi_\theta(\cdot|q)\|\pi_{\text{ref}}(\cdot|q)) \tag{1}$$

2.3 LoRA

Fine-tuning a full foundation model on a new task can be is computationally prohibitive. Hence, Parameter-Efficient Fine-Tuning (PEFT) methods have become standard. Among the most successful of these is Low-Rank Adaptation (LoRA) [14], which assumes that the change in a model's weights

during fine-tuning (ΔW), has a low intrinsic rank. Instead of learning ΔW directly, LoRA approximates it with the product of two much smaller matrices. We apply LoRA by configuring it with a rank $r = 16$ and setting the scaling factor $\alpha = 2 * r$, a common and effective heuristic.

2.4 Risk Management and Controls

Risk in finance is the possibility that an activity will have outcomes that differ from expectations, potentially resulting in financial loss or other adverse effects [2, 16]. Risk management in finance is essential to ensure that banks can identify, assess, and control risks to reduce the likelihood and impact of unexpected losses [1]. Reasons include complying with regulatory requirements (Basel III [3], PRA [4], FCA [12], etc.), ensuring business continuity during market shocks, and supporting strategic planning by allowing firms to better anticipate and respond to unexpected events. Additionally, strong risk management boosts the reputation of the firm and investor confidence [32, 15, 33].

Risk controls are all methods to reduce frequency and/or severity of uncertain events, this includes the prevention of risk and detection of risk [23]. There are 3 different types of controls: preventive controls (prevent risk), detective controls (identify risk) and corrective controls (applied upon discovery). They ensure risk exposures remain within a banks' risk appetite for effective risk management [33, 34].

3 Methodology

3.1 Dataset

We integrate the FinGPT-Forecaster dataset [37], which contains news headlines and financial data for the DJIA over a specific period [18]. For each entry in the base dataset, we calculated six critical risk and volatility metric [36, 21]s: VaR 1%, VaR 5%, CVaR 1%, CVaR 5%, Vol 7d, Vol 30d. We employ an 80/20 train test split. Finally, all evaluations are done on the test set (for all models).

3.1.1 Volatility Calculation

Volatility serves as a measure of the dispersion of returns for a given security. We calculated it over 7 days and 30 days. For a given series of closing prices $\{P_0, P_1, \ldots, P_n\}$, the daily log returns are:

$$R_t = \ln\left(\frac{P_t}{P_{t-1}}\right) \tag{2}$$

The volatility (StDev) of R_t is annualized to standardize comparisons (T = trading days in a year):

$$\sigma_{\text{annual}} = \text{StDev}(R_{t-N+1}, \ldots, R_t) \times \sqrt{T} \tag{3}$$

3.1.2 VaR and CVaR

To assess risk, we calculated the VaR and CVaR using a historical simulation approach, which leverages the distribution of past returns [36]. Given the historical daily returns $\{R_t\}$ over the 90-day window, the VaR at a confidence level of $1 - \alpha$ is the negative of the α-quantile of this empirical distribution. We calculated VaR and CVaR for two standard percentiles: 5% and 1%:

$$\text{VaR}_{1-\alpha}(\{R_t\}) = -\text{Quantile}(\{R_t\}, \alpha) \tag{4}$$

CVaR, or Expected Shortfall, describes the expected loss given that the loss exceeds the VaR. Let the set of returns in the tail of the distribution be $\mathcal{L} = \{R_t \mid R_t \leq -\text{VaR}_{1-\alpha}\}$. The CVaR is then:

$$\text{CVaR}_{1-\alpha}(\{R_t\}) = -\frac{1}{|\mathcal{L}|} \sum_{R_t \in \mathcal{L}} R_t \tag{5}$$

3.2 Training Risko1

Continuous metrics are discretized in bands (e.g. "0.2–0.3" for volatility), and the midpoint is used to compute the MSE. We fine-tune Fino1-8B using LoRA and GRPO with 3 rewards: formatting, selected band correctness (+100 exact, +10 adjacent, and -MSE otherwise), and constraints (CVaR $>$ VaR, higher confidence $\Rightarrow$ higher risk, non-negativity). Details for reproducibility in Appendix D.

Model	Constr.	Avg	Vol 7d	Vol 30d	VaR 1%	VaR 5%	CVaR 1%	CVaR 5%
General Models								
Meta-Llama-3.1-8B-Instruct	15.1	10.5	13.2	9.4	20.1	5.7	12.6	1.9
Meta-Llama-3.3-70B-Instruct	**100.0**	28.3	30.1	4.5	35.0	**58.1**	6.5	35.4
Specialised Models								
Fino1-8b	20.0	14.9	16.5	17.6	20.0	17.6	4.7	12.9
Fino1-14b	73.7	20.2	**36.8**	21.1	26.3	10.5	15.8	10.5
Risko1	99.1	**34.7**	20.5	**42.4**	**36.6**	45.1	**27.2**	**36.2**

Table 1: Accuracy (%) across volatility, VaR, and CVaR bands, with constraint compliance. The best performance in bold, and the second best in underline. **Takeaway: Risko1 achieves near-perfect constraint satisfaction and outperforms in accuracy specialized models and larger models.**

3.3 Classifying Generated Risks and Assigning Controls with the Taxonomy

Regulatory frameworks require financial institutions to establish structured processes for identifying and managing risks. In particular, banks need to develop and maintain a structured risk taxonomy: a classification (often hierarchical) of different risks an institution might face [5]. Typically, the process begins by identifying level 1 risks, such as market, credit, operational, and liquidity risks. Each can then be broken down into more specific level 2 risks. For instance, within operational risk, level 2 risks may include Information Technology (IT) and Information Security (IS) risks, legal risks, third-party risks, and fraud risks [26]. Further granularity can be achieved by defining level 3 risks; for example, under IT risk, level 3 risks might include unauthorized access and software failures. Additional sub-levels can be introduced for precise categorization and monitoring.

Controls should be assigned to the most granular level of risk within the taxonomy. If level 3 is the highest level of detail for a risk, then controls are mapped to the level 3 risk; if a more detailed level 4 risk is identified, controls are mapped at that level instead. For example, within unauthorized access (a level 3 risk), a level 4 risk might be terminated employee's unauthorized access. Controls addressing this could include IT off boarding and revocation of system access upon termination.

4 Results

We evaluate risk scenarios generated by Risko1 against the ground truth in terms of volatility at 7 and 30 days, and VaR/CVaR at the 1% and 5% confidence levels. We focus on the Llama family of models [13], as it supports comparisons across scale while keeping prompting constant, and it includes Fino1-8B and Fino1-14B, the best specialist models. We evaluate exact and tolerance band accuracy, MSE, and constraint/format compliance. All evaluations were carried out on the test set.

Results (Table 1) for Risko1 show **near-perfect constraint compliance**, and **higher accuracy than all other models** with **MSE comparable to Llama 3.3 70B and lower than all other models**, while using an order of magnitude fewer parameters. Accuracy (tolerance +- 1) and MSE in Appendix A.

4.1 Predictivity

We further explore the predictivity of Risko1, and check the relative performance of our model in the context of VaR/CVar/Volatility calculated in the time period after the current row in the dataset. We train a separate model for this, and we note that this stabilizes convergence. Results are similar to training on the historical method, and are reported in Appendix C.

5 Analysis: Assigning Controls

We first generate risk scenarios with Risko1 from financials and news in the test set, then use GPT-5 [27] to classify them under the standard risk taxonomy. Out of 725 generated risks, **over 88% were classifiable**. Categories were aligned with input sources 55% were Market (from financials) and 35% Operational (from news). The distribution shows strong granularity, with most risks being tier 3 (51%) of risks and tier 4 (20%), demonstrating that Risko1 produces precise, actionable risks suitable for targeted management. See Appendix B for complete charts.

6 Limitations and Future Directions

Unimodal LLMs are not as effective as multi-modal LLMs for risk prediction [6]. While we focus on scenario generation rather than prediction, this represents a future direction. Another limitation is scale, which we are working on as a future direction, with models an order of magnitude larger.

7 Conclusion

We introduce **Risko1**, a domain-tuned financial reasoning model that generates risk scenarios competitive with models one order of magnitude larger in terms of VaR, CVaR, and volatility under structural constraints. These results show that reinforcement-optimized, domain-specific training enables compact models to deliver reliable, interpretable risk assessments.

8 Acknowledgements

We acknowledge and thank Nscale for providing the compute resources (8 AMD Mi250x GPUs) used for all training and evaluations in this paper. We are especially grateful to Karl Havard for leading this partnership, Konstantinos Mouzakitis for his technical assistance, Brian Dervan for structuring our collaboration, and the entire Nscale team for their support.

References

[1] O. Adeyemi, A. Adeoye, and A. A. Adeleke. Integrating ai in financial risk management: Evaluating the effects of machine learning algorithms on predictive accuracy and regulatory compliance. *International Journal of Scientific Research and Analysis*, 13(2):6482–6489, 2024.

[2] T. Aven. Risk assessment and risk management: Review of recent advances on their foundation, Dec 2016.

[3] Bank for International Settlements. Basel iii: International regulatory framework for banks, Dec. 2017. Accessed: 2025-08-30.

[4] Bank of England Prudential Regulation Authority. Operational resilience: Impact tolerances for important business services, Mar. 2022. Accessed: 2025-08-29.

[5] B. Boultwood and J. Switzer. Erm: The importance of building a risk taxonomy, Apr. 2025. Accessed: 2025-08-30.

[6] Y. Cao, Z. Chen, P. Kumar, Q. Pei, Y. Yu, H. Li, F. Dimino, L. Ausiello, K. P. Subbalakshmi, and P. M. Ndiaye. Risklabs: Predicting financial risk using large language model based on multimodal and multi-sources data, 2025.

[7] H. Chen and J. Chen. Deepseek in Enterprise Intelligent Financial Alerting. *Scientific Journal of Economics and Management Research*, 2:8740–8747, 2024.

[8] M. Chen, M. DeHaven, I. Kitschelt, S. J. Lee, and M. J. Sicilian. Identifying Financial Crises Using Machine Learning on Textual Data. *Journal of Risk and Financial Management*, 16(3):161, 2023.

[9] Z. Chu, H. Guo, X. Zhou, Y. Wang, F. Yu, H. Chen, W. Xu, X. Lu, Q. Cui, L. Li, J. Zhou, and S. Li. Data-Centric Financial Large Language Models. *arXiv preprint arXiv:2310.17784*, 2023.

[10] DeepSeek-AI, D. Guo, D. Yang, H. Zhang, J. Song, R. Zhang, R. Xu, Q. Zhu, S. Ma, P. Wang, X. Bi, X. Zhang, X. Yu, Y. Wu, Z. F. Wu, Z. Gou, Z. Shao, Z. Li, Z. Gao, A. Liu, et al. Deepseek-r1: Incentivizing reasoning capability in llms via reinforcement learning, 2025.

[11] A. P. Desai, G. S. Mallya, M. Luqman, T. Ravi, N. Kota, and P. Yadav. Opportunities and challenges of generative-ai in finance, 2025.

[12] Financial Conduct Authority. Fg20/1: Assessing adequate financial resources, June 2020. Accessed: 2025-08-30.

[13] A. Grattafiori, A. Dubey, A. Jauhri, A. Pandey, A. Kadian, A. Al-Dahle, A. Letman, A. Mathur, A. Schelten, A. Vaughan, A. Yang, A. Fan, A. Goyal, A. Hartshorn, A. Yang, A. Mitra, A. Sravankumar, A. Korenev, A. Hinsvark, et al. The llama 3 herd of models, 2024.

[14] E. J. Hu, Y. Shen, P. Wallis, Z. Allen-Zhu, Y. Li, S. Wang, L. Wang, and W. Chen. Lora: Low-rank adaptation of large language models, 2021.

[15] J. Krause. Automated Repair of AI Code with Large Language Models and Formal Verification. *arXiv preprint arXiv:2405.08848*, 2024.

[16] S. S. S. Kumar and S. Sabahat. Emerging trends of risk management in the investment sector in india: A review paper. *EPRA International Journal of Multidisciplinary Research*, 10:15996–16005, 2024.

[17] X. Li. Large Language Models for Corporate Financial Distress Prediction: Overview and Exploration. *Highlights in Business, Economics and Management*, 47:471–477, 2024.

[18] J. Liu, Y. Wang, J. Li, Y. Li, H. Li, Y. Liu, and L. Chen. An applied study of multimodal financial data for financial early warning of listed companies. *IEEE Xplore*, 2024.

[19] X.-Y. Liu, X. Zhang, Q. Ding, Z. Liu, Y. Wang, H. Lei, S. Zheng, J. Guo, and C. D. Wang. Parallel market environments for finrl contests. In *2024 IEEE International Conference on Learning Representation in Finance (FinRL)*, pages 1–8, 2024.

[20] Z. Liu, X. Guo, F. Lou, L. Zeng, J. Niu, Z. Wang, J. Xu, W. Cai, Z. Yang, X. Zhao, et al. Fin-r1: A large language model for financial reasoning through reinforcement learning. *arXiv preprint arXiv:2503.16252*, 2025.

[21] Z. Liu and Y. Zhao. Deep learning models for price forecasting of financial time series: A review of recent advancements: 2020–2022. *WIREs Data Mining and Knowledge Discovery*, 13:e1519, 2023.

[22] L. Madmoni, A. Zait, I. Labzovsky, and D. Karmon. The Ability of Large Language Models to Evaluate Constraint-satisfaction in Agent Responses to Open-ended Requests. *arXiv preprint arXiv:2409.14371*, 2024.

[23] O. R. Manual. Risk controls — open risk manual,, 2020. [Online; accessed 30-August-2025].

[24] F. S. Mishkin. Over the cliff: From the subprime to the global financial crisis. *Journal of Economic Perspectives*, 25(1):49–70, March 2011.

[25] P. Nagy, S. Frey, K. Li, B. Sarkar, S. Vyetrenko, S. Zohren, A. Calinescu, and J. Foerster. Lob-bench: Benchmarking generative ai for finance – an application to limit order book data, 2025.

[26] Nomura Holdings, Inc. Risk management, 2025. Accessed: 2025-08-30.

[27] OpenAI. Gpt-5, 2025.

[28] L. Qian, W. Zhou, Y. Wang, X. Peng, J. Huang, and Q. Xie. Fin-01: On the transferability of reasoning-enhanced llms and reinforcement learning to finance. *arXiv preprint arXiv:2502.08127*, 2025.

[29] B. Saha, N. Rani, and S. K. Shukla. Generative ai in financial institution: A global survey of opportunities, threats, and regulation, 2025.

[30] R. Sawhney, P. Khanna, A. Aggarwal, T. Jain, P. Mathur, and R. R. Shah. VolTAGE: Volatility Forecasting via Text-Audio Fusion with Graph Convolution Networks for Earnings Calls. *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 643–653, 2020.

[31] Z. Shao, P. Wang, Q. Zhu, R. Xu, J. Song, X. Bi, H. Zhang, M. Zhang, Y. Li, Y. Wu, et al. Deepseekmath: Pushing the limits of mathematical reasoning in open language models. *arXiv preprint arXiv:2402.03300*, 2024.

[32] V. Shrivastava, J. Balasubramanian, A. Katyal, A. Yadav, and S. Yogananthan. Understanding the significance of risk management in enterprise management dynamics. *Multidisciplinary Reviews*, 6:2023ss093, 05 2024.

[33] A. Smith and S. Johnson. Automated compliance verification for AI models in enterprise Cloud MLOps Pipelines. *World Journal of Advanced Research and Reviews*, 21:2022–2030, 2024.

[34] H. Stokes. Sanctions and export control evasion techniques and practical steps to identify, assess and mitigate risk. `https://hstalks.com/doi/10.69554/DHAH5485/`, 2024.

[35] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin. Attention is all you need, 2023.

[36] Y. Wang and X. Xiang. FINANCIAL TIME SERIES FORECASTING USING ADAPTIVE RISK METRICS AND TRANSFORMER MODELS. *Upubscience Publisher*, 11:63–71, 2024.

[37] H. Yang, X.-Y. Liu, and C. D. Wang. Fingpt: Open-source financial large language models. *FinLLM Symposium at IJCAI 2023*, 2023.

[38] S. Yildirim, S. F. Günal, and A. Ayar. TULIP: Adapting Open-Source Large Language Models for Underrepresented Languages and Specialized Financial Tasks. *arXiv preprint arXiv:2407.03498*, 2024.

[39] L. Zhang. A review of enterprise management and risk assessment methods in the field of corporate finance in the past decade, Jan 2025.

[40] J. Černevičienė and A. Kabašinskas. Explainable artificial intelligence (xai) in finance: A systematic literature review, 2024.

Model	Vol 7d	Vol 30d	VaR 1%	VaR 5%	CVaR 1%	CVaR 5%
General Models						
Meta-Llama-3.1-8B-Instruct	0.084	0.165	5.157	16.7	13.6	38.3
Meta-Llama-3.3-70B-Instruct	0.017	0.061	**1.700**	**0.556**	7.327	**1.263**
Specialised Models						
Fino1-8b	0.074	0.127	5.940	12.8	13.5	21.9
Fino1-14b	0.020	0.026	23.8	11.8	53.4	26.3
Risko1	**0.017**	**0.009**	1.909	0.672	**4.792**	1.578

Table 2: Model Mean Squared Error (MSE) by Metric. The best performance in bold, and the second best in underline. **Takeaway: Risko1 and Llama 3.3 70B score comparably across risk metrics in terms of MSE.**

Model	Avg	Vol7d	Vol30d	VaR1	VaR5	CVaR1	CVaR5
General Models							
Meta-Llama-3.1-8B-Instruct	31.3	45.3	30.8	47.8	24.5	32.1	7.5
Meta-Llama-3.3-70B-Instruct	65.2	**82.1**	24.4	76.4	93.9	29.3	**85.4**
Specialised Models							
Fino1-8b	44.5	56.5	54.1	55.3	37.6	34.1	29.4
Fino1-14b	44.7	68.4	52.6	47.4	52.6	31.6	15.8
Risko1	**82.4**	78.1	**92.0**	**79.9**	**95.1**	**67.9**	81.2

Table 3: Model Accuracy (%) by Metric - Within One Band. The best performance in bold, and the second best in underline. **Takeaway: Risko1 performs best, with Llama 3.3 70B a close second.**

A Full Model Comparison Results

We report full metrics for MSE and tolerance accuracy for both Risko1, our baselines in the Llama herd of models, and specialized LLMs in Table 2 and Table 3.

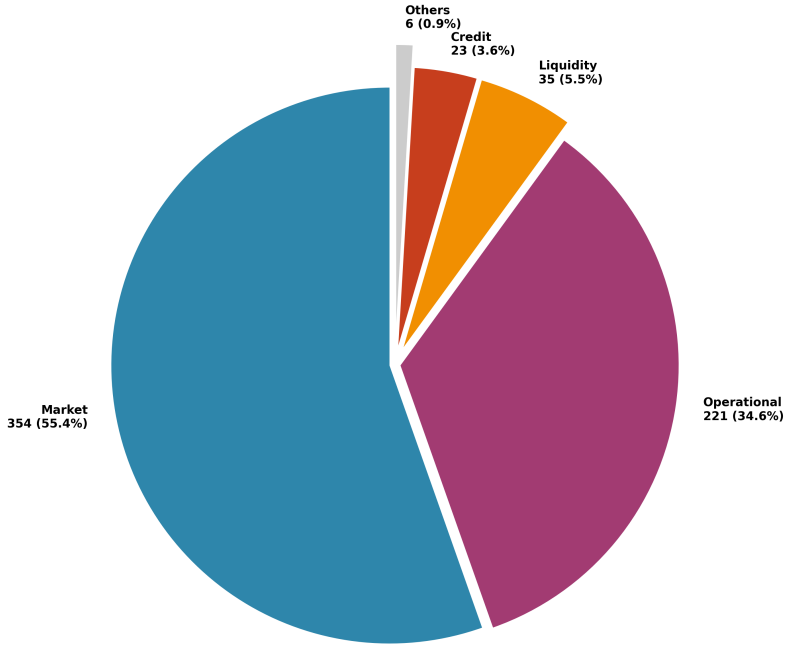

Figure 2: Share of the categories of the risks generated by Risko1. **Takeaway: risks are generated across categories of risks, and reflect the input data Risko1 received: with market risks from financials and operational risks coming from news.**

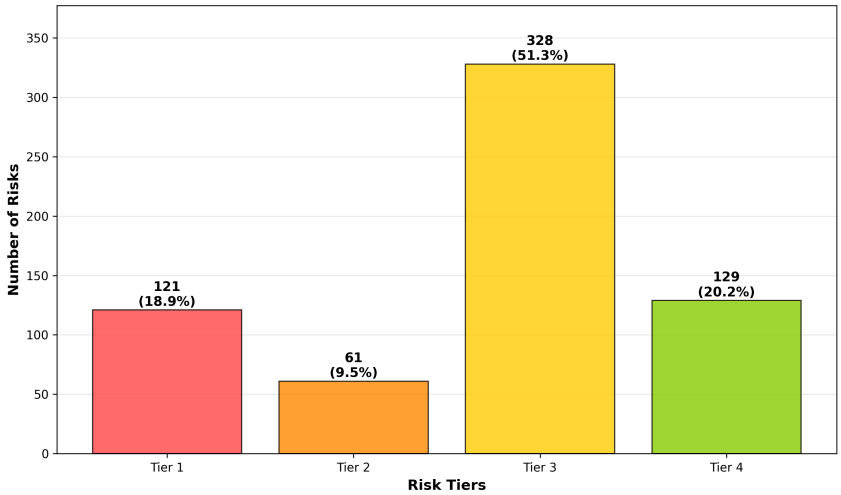

Figure 3: Distribution of the most granular tier for the risks generated by Risko1 in the taxonomy. **Takeaway: generated risks are granular, with most of them falling in tiers 3 and 4. This allows for specific controls to be put in place.**

B Full Analysis Results

We report figures for our analysis of risks scenarios generated by Risko1, in particular Figure 2 chart of categories of the risks we generate, and in Figure 3 we report the distributions of the most granular tiers our risks can be classified in.

Model	Constr.	Avg	Vol 7d	Vol 30d	VaR 1%	VaR 5%	CVaR 1%	CVaR 5%
General Models								
Meta-Llama-3.1-8B-Instruct	17.2	10.5	17.2	11.9	19.2	6.0	6.6	2.0
Meta-Llama-3.3-70B-Instruct	**100.0**	26.7	**27.2**	4.1	30.9	**56.5**	7.3	**34.1**
Specialised Models								
Fino1-14b	71.8	15.0	17.9	15.4	20.5	20.5	10.3	5.1
Fino1-8b	22.7	16.9	17.0	20.5	29.5	12.5	13.6	8.0
Risko1	98.3	**33.5**	22.7	**37.8**	**36.6**	44.1	**29.8**	29.8

Table 4: Accuracy (%, future) across volatility, VaR, and CVaR bands, with constraint compliance. The best performance in bold, and the second best in underline. **Takeaway: Risko1 achieves near-perfect constraint satisfaction and outperforms specialized models and larger models.**

Model	Avg	Vol7d	Vol30d	VaR1	VaR5	CVaR1	CVaR5
General Models							
Meta-Llama-3.1-8B-Instruct	35.0	49.0	30.5	59.6	25.8	33.8	11.3
Meta-Llama-3.3-70B-Instruct	65.7	**82.1**	24.0	74.4	**96.3**	32.5	**85.0**
Specialised Models							
Fino1-14b	52.6	66.7	56.4	59.0	56.4	38.5	38.5
Fino1-8b	42.8	51.1	43.2	56.8	38.6	37.5	29.5
Risko1	**79.8**	73.1	**87.0**	**78.6**	92.4	**71.4**	76.5

Table 5: Model Accuracy (%, future) by Metric - Within One Band. Best performance in bold, and second best underlined. **Takeaway: Risko1 performs best, with Llama 3.3 70B a close second.**

Model	Vol 7d	Vol 30d	VaR 1%	VaR 5%	CVaR 1%	CVaR 5%
General Models						
Meta-Llama-3.1-8B-Instruct	0.085	0.228	6.086	16.5	16.1	40.4
Meta-Llama-3.3-70B-Instruct	**0.015**	0.055	**1.948**	**0.530**	**7.434**	**1.356**
Specialised Models						
Fino1-14b	0.038	0.045	6.860	6.106	16.7	11.8
Fino1-8b	0.069	0.100	5.350	9.872	13.8	18.3
Risko1	0.032	**0.017**	3.167	1.742	7.738	3.692

Table 6: Model Mean Squared Error (MSE, future) by Metric. The best performance in bold, and the second best in underline. **Takeaway: Risko1 and Llama 3.3 70B score similarly.**

C Predictivity

D Reproducibility Details

This appendix provides comprehensive implementation details for reproducing all training, evaluations, and experiments in this paper.

D.1 Data

D.1.1 Dataset Description

The dataset consists of financial news and context data for DJIA constituents, sourced from FinGPT-Forecaster. The dataset contains 1215 total samples distributed as follows:

- Training set: 972 samples (80%)
- Test set: 243 samples (20%)

Each sample includes:

- Company symbol (DJIA constituent)
- Time period (date range format: YYYY-MM-DD to YYYY-MM-DD)
- Financial news context and company information
- Ground truth risk metrics: volatility_7d, volatility_30d, var_1p, var_5p, cvar_1p, cvar_5p (withheld from model during training and used for reward calculation and evaluation)

D.1.2 Preprocessing Pipeline

Returns Calculation: VaR and CVaR values are converted from negative decimal format to positive percentage format using the transformation:

$$\text{percentage} = |\text{decimal_value}| \times 100 \tag{6}$$

Volatility Processing: Volatility metrics (7-day and 30-day) are used directly in decimal format without transformation.

Data Sanitization: Legacy instruction wrappers are removed from prompts using regex patterns to strip `[INST]`, `[/INST]`, `«SYS»`, and `«/SYS»` tags.

D.1.3 Label Construction

Risk Metrics Definitions:

- **VaR (Value at Risk):** 1% and 5% confidence levels
- **CVaR (Conditional Value at Risk):** 1% and 5% confidence levels
- **Volatility:** 7-day and 30-day rolling windows

Confidence Levels: VaR and CVaR computed at 1% and 5% confidence levels, with fundamental constraints:

- CVaR > VaR at any confidence level
- 1% confidence metrics > 5% confidence metrics

D.1.4 Discretization Scheme

Percentage Bands (VaR/CVaR): 18 bands with variable widths:

```
"0-1%", "1-2%", "2-3%", "3-4%", "4-5%", "5-6%", "6-7%",
"7-8%", "8-9%", "9-10%", "10-12%", "12-15%", "15-20%",
"20-25%", "25-30%", "30-40%", "40-50%", "50%+"
```

Volatility Bands: 14 bands with variable widths:

```
"0-0.1", "0.1-0.2", "0.2-0.3", "0.3-0.4", "0.4-0.5",
"0.5-0.6", "0.6-0.7", "0.7-0.8", "0.8-0.9", "0.9-1.0",
"1.0-1.2", "1.2-1.5", "1.5-2.0", "2.0+"
```

Band widths increase for higher values to account for increased uncertainty and reduced precision requirements at extreme values.

D.1.5 Data Splits

Split Method: Random 80/20 train-test split using `seed=42` via HuggingFace datasets `train_test_split()`.

D.1.6 Data Cleaning

Missing Values: Converted to NaN using robust `to_float()` function with exception handling.

Outlier Treatment: No explicit outlier removal; extreme values handled through open-ended bands ("50%+", "2.0+").

Corporate Actions: No explicit handling of splits or dividends in preprocessing.

D.2 Model Architecture

D.2.1 Base Model

Architecture: TheFinAI/Fino1-8B (Llama 3.1-8B based financial model)

- Parameter count: 8 billion parameters
- Vocabulary size: 128,256 tokens (Llama 3.1 tokenizer)
- Context length: 512 tokens (prompt) + 1024 tokens (completion)
- Precision: bfloat16 for training

Tokenizer: Llama 3.1 tokenizer with:

- Padding side: right
- Pad token: EOS token (128001)
- Special tokens preserved

D.2.2 LoRA Configuration

Target Modules: Default PEFT auto-detection for Llama (typically q_proj, k_proj, v_proj, o_proj, gate_proj, up_proj, down_proj)

LoRA Parameters:

- Rank (r): 16
- Alpha: 32
- Dropout: 0.05
- Bias: none
- Task type: CAUSAL_LM

D.3 Training Details

D.3.1 Supervised Fine-tuning

No supervised fine-tuning phase before GRPO. Training begins directly with GRPO from the pre-trained TheFinAI/Fino1-8B checkpoint.

D.3.2 GRPO Setup

Group Parameters:

- Rollouts per prompt: 7 generations
- KL penalty coefficient (β): 0 (pure reward optimization)
- Sampling temperature: 1.0 (default vLLM)
- Top-p: 1.0 (no nucleus sampling)

Reward Function Weights: Equal weighting (1.0 each) for three reward components:

- **Formatting reward:** +10 for valid JSON with correct schema, 0 otherwise. +3 if the reasoning is between 500 and 700 tokens (to encourage detailed explanations whithin VRAM constraints)
- **Value reward:** Band-based scoring: +100 for exact band match, +10 for adjacent band, -MSE(gt_value, band_midpoint) for distant bands
- **Constraint reward:** +50 for satisfying all risk metric ordering constraints, 0 otherwise

Training Horizon:

- Training steps: Until convergence (checkpoint-based resumption)
- Save frequency: Every 50 steps
- Logging frequency: Every 10 steps

Batch Configuration:

- Per-device batch size: 1
- Gradient accumulation steps: 1
- Effective batch size: 1×7 generations = 7

D.3.3 Optimizer

Optimizer Type: AdamW

Learning Rate Schedule:

- Learning rate: 1e-5
- Schedule: Decay
- Weight decay: linear decay to 0
- Betas: Default (0.9, 0.999)
- Epsilon: Default (1e-8)

D.3.4 Compute Setup

Hardware Configuration:

- GPU Type: AMD Mi250x GPUs with ROCm 6.3.1
- GPU Count: 8 GPUs per node
- Memory: 64GB VRAM per GPU
- Tensor Parallel Size: 8 (for inference)
- Docker Environment: ROCm/vLLM:rocm6.3.1_vllm_0.8.5_20250513

Training Infrastructure:

- Framework: TRL 0.16.1
- Backend: vLLM for generation, transformers for training

- Mixed Precision: bfloat16
- Distribution: Single-node multi-GPU

Environment Variables:

```
TORCH_DISTRIBUTED_USE_DTENSOR=0
TORCH_DIST_DDP_SHARDING=0
ACCELERATE_USE_TP=false
PYTORCH_ENABLE_DISTRIBUTED=1
PYTORCH_CUDA_ALLOC_CONF=expandable_segments:True
TORCH_COMPILE_DISABLE=1
NCCL_IB_DISABLE=1
HSA_ENABLE_SDMA=0
```

D.4 Evaluation Protocol

D.4.1 Metrics

Band Classification Metrics:

- **Exact Accuracy:** Fraction of predictions matching ground truth band exactly
- **Tolerance Accuracy:** Fraction within ±1 band of ground truth
- **Macro-averaged Precision/Recall/F1:** Multi-class metrics across all bands
- **MSE:** Mean squared error between band midpoints and ground truth values

Format Compliance:

- JSON validity rate
- Schema compliance rate (correct keys and data types)

Constraint Compliance: Percentage of predictions satisfying all four ordering constraints:

- CVaR_1p > VaR_1p
- CVaR_5p > VaR_5p
- VaR_1p > VaR_5p
- CVaR_1p > CVaR_5p

D.4.2 Scenario Taxonomy Evaluation

GPT-5 Classification Pipeline:

- Model: OpenAI GPT-5 via API
- Concurrency: 8 parallel requests
- Classification Scope: Only risks identified by the model (no new risks added)

Risk Taxonomy Hierarchy:

- **Tier 1:** Market, Credit, Operational, Liquidity
- **Tier 2:** IT/IS risks, legal risks, third-party risks, transaction lifecycle, fraud
- **Tier 3:** Data breaches, unauthorized access, software failures
- **Tier 4:** Specific instances (e.g., terminated employee access)

Sampling Parameters:

- Number of scenarios: All test set samples with valid risk predictions
- Risk extraction: From "risks" field in model JSON output

- Unclassifiable handling: Separate category for vague/generic items

GPT-5 Prompt:

```
You are a risk analyst.
You will be presented with:
- company financial information
- news about the company
- short term risk metrics according to another analyst
- reasoning behind the metrics, and specific risks identified

Follow this taxonomy when classifying risks:

A risk taxonomy is a structured classification (often hierarchical)
of different risks an institution might face.
Level 1: market, credit, operational, liquidity.
Level 2: more specific, for example, within operational, typical level 2
include Information Technology (IT) and Information Security (IS) risks,
legal risks, third-party risks, transaction lifecycle risks, and fraud risks.
Level 3: provides additional granularity; for example, within IT risk,
level 3 might list data breaches, unauthorised access, and software failures.
Level 4: even more granular risks (eg unauthorised access by a terminated employee).

Controls must be proportional:
- Map each control to the most granular level in the respective risk's taxonomy.
- Avoid both under-tagging (assigning too broad a tier) and over-tagging
  (assigning too narrow a tier), but focus on what is appropriate for the
  specific risk being assessed.

For items that cannot be classified as risks or are too vague/generic to map
to the taxonomy, use:
- taxonomy_path: "Unclassifiable"
- tier: "N/A"
- control: "Not a specific risk" or explain why it cannot be classified

Your answer format must be valid JSON with keys:
{
  "identified_risks": [
    {
      "taxonomy_path": "Operational - IT - Unauthorised Access - Terminated Employee",
      "tier": "Tier 4",
      "control": "Immediate revocation of credentials upon termination and an
                  HR-IT offboarding checklist"
    },
    {
      "taxonomy_path": "Unclassifiable",
      "tier": "N/A",
      "control": "Not a specific risk - too generic/vague to map to operational
                  taxonomy"
    }
  ]
}

Company: {symbol}

Company Context and News:
{context[:2000]}...

Risk Analyst Prediction and Reasoning:
{generated_text[:1500]}...
```

```
Identified Risks by Previous Analyst:
{', '.join(risks)}

Please classify exactly and only these risks (don't add new risks) using
the risk taxonomy and assign appropriate tier levels (Tier 1-4) based on
granularity. Also suggest appropriate controls.

For each risk, provide:
1. The appropriate taxonomy path
2. The appropriate tier level (Tier 1-4)
3. Controls
```

D.5 Baselines

D.5.1 Models Compared

Model Versions:

- **Meta-Llama-3.1-8B-Instruct:** Base instruction-tuned model
- **Meta-Llama-3.3-70B-Instruct:** Latest instruction-tuned variant
- **TheFinAI/Fino1-8B:** Base financial model (before GRPO)
- **TheFinAI/Fin-o1-14B:** Larger financial reasoning model
- **Risko1:** Our GRPO-trained model

Access Mode: All models served locally via vLLM with identical infrastructure.

Sampling Parameters:

- Temperature: 1.0
- Top-p: 1.0
- Max tokens: 1024
- Stop tokens: Default EOS handling

D.5.2 Evaluation Consistency

Data Consistency:

- Identical test split (seed=42, 80/20% split)
- Same band definitions and boundaries
- Identical constraint checking logic
- Same prompt templates and system instructions

Reproducibility Notes:

- GRPO training involves stochastic generation; results may vary between runs
- Data splits and evaluation metrics are deterministic given fixed seeds
- GPT-5 risk classification introduces additional non-determinism
- Cached completions provided for exact result reproduction
- Hardware-specific optimizations may affect training dynamics