Proceedings of Machine Learning Research **318** (2026)
The 39th Canadian Conference on Artificial Intelligence
Canadian AI 2026
25-29 May 2026,Vancouver,BC

STRIDE Moves Market Sentiment

Sujay Uday Rittikar[†,*], Sheela Ramanna[†]
[†] Applied Computer Science, The University of Winnipeg

Abstract

Aspect-based sentiment analysis in the financial domain requires models to reason over sparse, entity-centric signals while remaining robust to linguistic variability and conflicting cues. We introduce **STRIDE**, a reinforcement learning framework that reformulates keyword selection as a sequential decision-making problem, integrating Directional Stimulus Prompting (DSP) with stable reward-driven policy optimization. To address the instability of sparse, high-variance reward signals, STRIDE incorporates exponential moving average (EMA) smoothing into the REINFORCE objective, enabling more reliable gradient estimates for policy learning. We evaluate STRIDE on two benchmark financial sentiment datasets: SEntFiN 1.0 and FinEntity. On SEntFiN 1.0, STRIDE achieves state-of-the-art F1-score (0.946) and near state-of-the-art accuracy (0.950). On FinEntity, STRIDE exceeds the previous state-of-the-art F1-score by 4.2%, achieving state-of-the-art performance on both accuracy (0.942) and F1-score (0.933). Across both datasets, the results demonstrate that EMA-smoothed rewards provide consistent improvements of 2.6% to 4.3% F1 relative to unsmoothed baselines, validating the effectiveness of stability-aware reward formulation for financial aspect-based sentiment analysis. The source code for reproducibility is available at: https://github.com/sujayrittikar/stride_sentiment_analysis.

Keywords: Reinforcement Learning, Financial Sentiment Analysis, Reward functions, Prompt Engineering

1. Introduction

The digital economy has transformed how market participants extract investment signals from social media and news discourse. Empirical evidence demonstrates this relevance: Bollen et al. [1] showed that Twitter mood predicts Dow Jones movements, while Ranco et al. [2] documented relationships between social media sentiment and abnormal returns, particularly in retail-dominated segments. However, extracting reliable signals presents substantial challenges. Market participants employ heterogeneous linguistic expressions including sarcasm, domain-specific terminology, and implicit framing, that complicate sentiment extraction when multiple aspects coexist within single documents. Recent advances in pretrained language models and aspect-based sentiment analysis have demonstrated promise [3], yet existing methods remain limited in adaptively selecting which textual signals are most informative for classification tasks [4]. While reinforcement learning has succeeded in optimizing NLP components, its application to financial aspect-based sentiment analysis remains underexplored. We propose STRIDE (Stimulus-guided Training for Reward Integration with Directional Exponential smoothing), a reinforcement learning framework grounded in fundamental policy gradient optimization. Our approach demonstrates that principled simplicity drives performance: through three key contributions, (1) adapting Direct Stimulus Prompting to aspect sentiment classification, (2) introducing Exponential Moving Average smoothing for stable reward estimation in REINFORCE, and (3) careful engineering of foundational principles, we achieve state-of-the-art results on SentFin 1.0 and FinEntity benchmarks.

[*] rittikar-s@webmail.uwinnipeg.ca

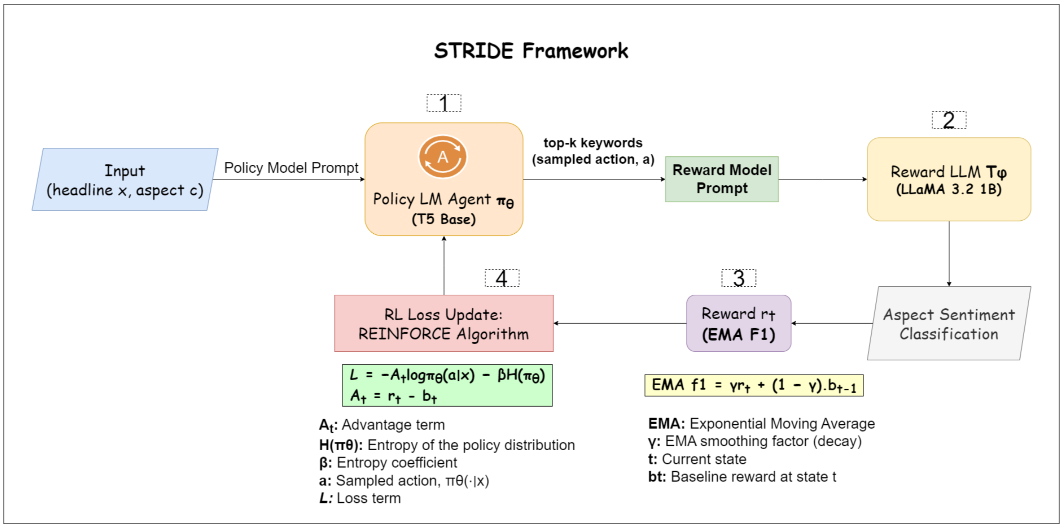

Figure 1. Overview of the REINFORCE + DSP framework for financial aspect-based sentiment analysis. The policy network learns to select task-relevant keywords that guide sentiment classification, with rewards smoothed via exponential moving averages to stabilize training.

2. Related Works

Benchmarking on financial sentiment datasets shows that domain-specific pre-training and structured architectural enhancements drive performance gains. On SEntFiN 1.0 [5], RoBERTa and FinBERT achieve strong baselines (94.29% accuracy, 93.27% F1), while on FinEntity [6], FinBERT-CRF establishes CRF-augmented architectures as effective for entity-level sentiment extraction. A more recent self-aware in-context learning approach [7] yields further gains via adaptive correction. However, these works focus on static architectural choices, leaving instance-adaptive signal selection largely unexplored. From adjacent NLP domains, EMA offers proven variance reduction and convergence stabilization in reinforcement learning [8], and DSP [9] demonstrates that instance-specific prompt optimization via policy gradients can yield 4–13% improvements in summarization, yet its application to financial sentiment remains unexplored.

3. Methodology

3.1. Datasets

Class	#Entities	Proportion
Neutral	5,517	38.3%
Positive	5,074	35.2%
Negative	3,813	26.5%

(a) SEntFiN 1.0 (14,404 entities)

Class	#Entities	Proportion
Neutral	1,048	51.9%
Positive	490	24.2%
Negative	483	23.9%

(b) FinEntity (2,021 entities)

Table 1. Sentiment class distribution across benchmark datasets.

We evaluate on two financial ABSA benchmarks. **SEntFiN 1.0** contains 10,753 financial news headlines with 14,404 entity–sentiment annotations, notable for its multi-entity complexity: 2,847 headlines cover multiple entities, of which 1,233 exhibit conflicting sentiments (Table 1a). **FinEntity** comprises 979 Reuters paragraphs with 2,021 annotated entities spanning equities, cryptocurrencies, and commodities, with 60% of texts containing entities of opposing polarity (Table 1b). Both datasets use a three-class taxonomy (positive, negative, neutral) with 80:10:10 train–validation–test splits.

3.2. STRIDE Framework

STRIDE is a reinforcement learning framework for financial ABSA comprising three components: a policy agent that selects task-relevant keywords, a frozen reward evaluator that scores sentiment quality, and an EMA-based reward formulation for training stability. The policy iteratively generates keyword stimuli, receives reward feedback, and updates via policy gradients.

3.2.1. Policy Agent and Reward Evaluator

The **Policy Agent** (π_θ) uses T5 Base [10], whose encoder–decoder architecture enables conditional keyword selection given a headline and target aspect, optimized via REINFORCE. The **Reward Evaluator** (T_ϕ) is a frozen LLaMA 3.2 1B [11] that classifies aspect-level sentiment on keyword-augmented inputs and returns F1-based rewards against gold labels.

3.2.2. Policy Optimization with REINFORCE

Given headline x and aspect c, the policy samples top-k keywords injected into the evaluator prompt via DSP, yielding reward r_t. The REINFORCE objective [12] with entropy regularization is:

$$L = -A_t \log \pi_\theta(a|x) - \beta H(\pi_\theta), \tag{3.1}$$

where $A_t = r_t - b_t$ is the advantage, b_t a baseline, $H(\pi_\theta)$ policy entropy, and β the exploration coefficient.

3.2.3. Directional Stimulus Prompting

DSP integrates the policy's keyword selections directly into the evaluator's input prompt as explicit stimuli for aspect-focused reasoning, tightly coupling keyword selection with downstream sentiment quality.

3.2.4. Reward Signal Formulation

STRIDE supports two reward variants. The *F1-only* setting uses raw F1 directly; the *EMA-smoothed* setting applies:

$$f1_t = \gamma r_t + (1-\gamma) f1_{t-1}, \tag{3.2}$$

where γ is the decay parameter, reducing gradient variance in policy updates.

3.2.5. Training Protocol

For each headline–aspect pair, the policy samples keywords injected via DSP into the evaluator prompt. The evaluator returns an F1-based reward, EMA-smoothed and used to compute the REINFORCE loss (Figure 1). Parameters are updated via gradient descent, jointly optimizing keyword selection and sentiment prediction in a modular, stable loop.

SEntFiN 1.0				FinEntity			
Methodology	**Acc.**	**F1**	**Mac-F1**	**Methodology**	**Acc.**	**F1**	**Mac-F1**
GPT-3.5 HAD [13]	0.782	0.777	—	GPT-3.5 MSV [13]	0.685	—	—
FinBERT [5]	0.911	0.933	—	FinBERT-CRF [6]	—	0.840	0.850
RoBERTa [5]	0.944	0.914	—	SILC [7]	—	0.891	—
DeBERTa+CLoRA [14]	**0.954**	0.944	—				
DSP-REINFORCE (ours)	0.895	0.903	0.893	DSP-REINFORCE (ours)	0.920	0.907	0.905
STRIDE (ours)	0.950	**0.946**	**0.935**	**STRIDE (ours)**	**0.942**	**0.933**	**0.923**

Table 2. Performance comparison on SEntFiN 1.0 (left) and FinEntity (right).

4. Results and Discussion

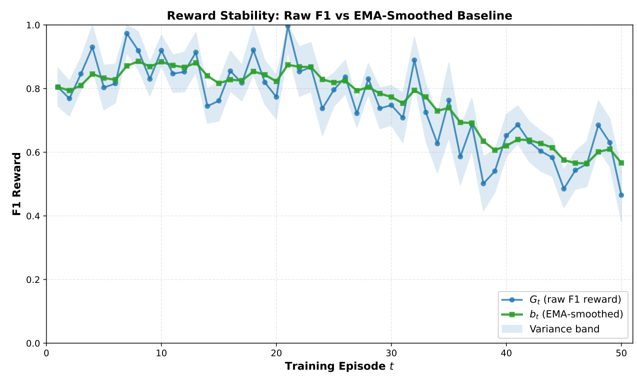

Figure 2. Reward stability on the SEntFiN 1.0 dataset during STRIDE training. The plot compares raw F1 rewards (blue line) exhibiting high variance against EMA-smoothed baseline estimates (green line). The shaded variance band illustrates the fluctuation in raw rewards, while the smoothed baseline provides stable gradient estimates for policy optimization.

Across both benchmarks (Table 2), **STRIDE** attains 0.950/0.946 accuracy/F1 on **SEntFiN 1.0** (+4.3% over the unsmoothed ablation) and 0.942/0.933 on **FinEntity** (+2.6%), with high Macro-F1 confirming balanced class-wise performance. As shown in Figure 2, EMA smoothing reduces reward variance and stabilizes REINFORCE updates. Prompts, class-wise breakdowns, and keyword analysis are provided in our github repository (see Abstract).

5. Conclusion

We introduced **STRIDE**, a REINFORCE-based framework for financial ABSA that achieves state-of-the-art F1 on FinEntity (+4.2%) and SEntFiN 1.0 (0.946) by reformulating keyword selection as sequential decision-making with EMA-stabilized rewards. Future work includes DPO [15] and KTO [16] for stimulus preference learning, and extension to cross-domain and multilingual settings [17].

Acknowledgments

This research is supported by NSERC Discovery grant #194376 and the University of Winnipeg President's Scholarship for World Leaders.

References

[1] J. Bollen, H. Mao, and X. Zeng. “Twitter mood predicts the stock market”. In: *Journal of Computational Science* 2.1 (Mar. 2011), pp. 1–8. ISSN: 1877-7503. DOI: 10.1016/j.jocs.2010.12.007.

[2] G. Ranco, D. Aleksovski, G. Caldarelli, M. Grčar, and I. Mozetič. “The Effects of Twitter Sentiment on Stock Price Returns”. In: *PLOS ONE* 10.9 (Sept. 2015). Ed. by T. Preis, e0138441. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0138441.

[3] K. Du, F. Xing, R. Mao, and E. Cambria. “Financial Sentiment Analysis: Techniques and Applications”. In: *ACM Computing Surveys* 56.9 (Apr. 2024), pp. 1–42. ISSN: 1557-7341. DOI: 10.1145/3649451.

[4] L. Shi, Z. Jia, and R. Liu. “Expert-level financial sentiment analysis via internal and external financial knowledge fusion”. In: *Information Fusion* 127 (Mar. 2026), p. 103884. ISSN: 1566-2535. DOI: 10.1016/j.inffus.2025.103884.

[5] A. Sinha, S. Kedas, R. Kumar, and P. Malo. “SEntFiN 1.0: Entity-aware sentiment analysis for financial news”. In: *Journal of the Association for Information Science and Technology* 73.9 (Mar. 2022), pp. 1314–1335. ISSN: 2330-1643. DOI: 10.1002/asi.24634.

[6] Y. Tang, Y. Yang, A. Huang, A. Tam, and J. Tang. “FinEntity: Entity-level Sentiment Classification for Financial Texts”. In: *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, 2023. DOI: 10.18653/v1/2023.emnlp-main.956.

[7] S. Zhu, C. He, H. Liu, P. Dong, H. Zhao, Y. Yan, Y. Jia, H. Zan, and M. Peng. “SILC-EFSA: Self-aware In-context Learning Correction for Entity-level Financial Sentiment Analysis”. In: *Proceedings of the 31st International Conference on Computational Linguistics*. Ed. by O. Rambow, L. Wanner, M. Apidianaki, H. Al-Khalifa, B. D. Eugenio, and S. Schockaert. Abu Dhabi, UAE: Association for Computational Linguistics, Jan. 2025, pp. 4980–4992. URL: https://aclanthology.org/2025.coling-main.333/.

[8] M. D. Awheda and H. M. Schwartz. “Exponential moving average based multiagent reinforcement learning algorithms”. In: *Artificial Intelligence Review* 45.3 (Oct. 2015), pp. 299–332. ISSN: 1573-7462. DOI: 10.1007/s10462-015-9447-5.

[9] Z. Li, B. Peng, P. He, M. Galley, J. Gao, and X. Yan. “Guiding Large Language Models via Directional Stimulus Prompting”. In: *Advances in Neural Information Processing Systems*. Ed. by A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine. Vol. 36. Curran Associates, Inc., 2023, pp. 62630–62656. URL: https://proceedings.neurips.cc/paper_files/paper/2023/file/c5601d99ed028448f29d1dae2e4a926d-Paper-Conference.pdf.

[10] C. Raffel, N. Shazeer, A. Roberts, K. Lee, S. Narang, M. Matena, Y. Zhou, W. Li, and P. J. Liu. “Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer”. In: *Journal of Machine Learning Research* 21.140 (2020), pp. 1–67. URL: http://jmlr.org/papers/v21/20-074.html.

[11] A. Dubey, A. Jauhri, A. Pandey, A. Kadian, A. Al-Dahle, A. Letman, A. Mathur, A. Schelten, A. Yang, A. Fan, et al. “The llama 3 herd of models”. In: *arXiv preprint arXiv:2407.21783* (2024). DOI: 10.48550/arXiv.2407.21783.

[12] M. Sewak. “Policy-Based Reinforcement Learning Approaches: Stochastic Policy Gradient and the REINFORCE Algorithm”. In: *Deep Reinforcement Learning*. Springer Singapore, 2019, pp. 127–140. ISBN: 9789811382857. DOI: 10.1007/978-981-13-8285-7_10.

[13] F. Xing. “Designing Heterogeneous LLM Agents for Financial Sentiment Analysis”. In: *ACM Transactions on Management Information Systems* 16.1 (Feb. 2025), pp. 1–24. ISSN: 2158-6578. DOI: 10.1145/3688399.

[14] V. Ivanenko. “Enhancing aspect-based financial sentiment analysis through contrastive learning”. In: *INNOVATIVE TECHNOLOGIES AND SCIENTIFIC SOLUTIONS FOR INDUSTRIES* 3(25) (Sept. 2023), pp. 138–147. ISSN: 2522-9818. DOI: 10.30837/itssi.2023.25.138.

[15] R. Rafailov, A. Sharma, E. Mitchell, C. D. Manning, S. Ermon, and C. Finn. “Direct Preference Optimization: Your Language Model is Secretly a Reward Model”. In: *Thirty-seventh Conference on Neural Information Processing Systems*. 2023. URL: https://openreview.net/forum?id=HPuSIXJaa9.

[16] K. Ethayarajh, W. Xu, N. Muennighoff, D. Jurafsky, and D. Kiela. “KTO: Model Alignment as Prospect Theoretic Optimization”. In: (Feb. 2024). DOI: 10.48550/ARXIV.2402.01306. arXiv: 2402.01306 [cs.LG].

[17] C. Wu, B. Ma, Z. Zhang, N. Deng, Y. He, and Y. Xue. “Evaluating zero-shot multilingual aspect-based sentiment analysis with large language models”. In: *International Journal of Machine Learning and Cybernetics* 16.10 (June 2025), pp. 8079–8101. ISSN: 1868-808X. DOI: 10.1007/s13042-025-02711-z.