WHEN RNN-BASED MARKED POINT PROCESSES FAIL IN REAL-WORLD FINANCE: A TINY PAPER

Anonymous authors
Paper under double-blind review

ABSTRACT

Neural Marked Temporal Point Process (MTPP) models have shown promise in controlled benchmarks for forecasting and event pattern modeling in finance. However, when deploying Recurrent Neural Network (RNN)-based MTPPs on large-scale, high-dimensional financial event streams, we encountered unexpected challenges: ballooning parameter sizes, increased computational costs, and training instability. This short paper outlines (1) the financial use case, (2) the literature-proposed neural MTPP solution, (3) the negative outcomes observed, and (4) our investigation into why standard MTPPs fail to generalize as promised in real-world conditions.

1 USE CASE: HIGH-DIMENSIONAL FINANCIAL EVENT STREAMS

Financial firms often track event streams at high frequency (e.g., limit-order book (LOB) updates or requests-for-quotes) for many instruments. Practitioners rely on forecast models for trading strategies, liquidity management, or risk assessment. Each event has a time stamp and a mark (e.g., transaction type), and these marks easily span dozens of categories.

2 SOLUTION PROPOSED IN DEEP LEARNING LITERATURE

Several works have proposed RNN-based MTPPs to capture cross-excitation among multiple marks in continuous time (Du et al., 2016; Mei & Eisner, 2017). The approach encodes past events in a hidden state and uses neural layers to parameterize the conditional intensity functions for different event types. On synthetic or low-dimensional benchmarks, these neural MTPPs have outperformed classical (e.g., Hawkes) processes.

3 NEGATIVE OUTCOMES IN REAL-WORLD DEPLOYMENTS

Parameter Explosion. When the number of event types or exogenous features exceeded a few dozen, the RNN input-to-hidden transformations became prohibitively large, leading to:

- *Memory and compute overhead*: Even moderate hidden sizes inflated total parameters into tens or hundreds of thousands, slowing training.
- *Overfitting and instability*: The network fit training data well but generalized poorly out-of-sample.

Domain Mismatch. We also found that standard neural MTPPs often assume data are "clean" and well-conditioned. In finance, microstructure noise, market closures, and heavy-tailed distributions break these assumptions, yielding inconsistent performance gains.

4 WHY IT DID NOT WORK AS EXPECTED

(1) Complex Input Structure. Contrary to simplified benchmarks, real financial features are not a single fixed-size vector; they arise from multi-asset, multi-venue, and multi-event interactions. Classic RNN layers scale poorly, as each new feature dimension inflates the weight matrices.

(2) Lack of Regularization for Cross-Excitation. While the RNN-based MTPP is powerful in principle, it lacks strong inductive biases for how real market events excite or dampen future events. Absent domain-inspired constraints, the model overfits ephemeral correlations.

(3) Computational and Latency Constraints. Finance often imposes tight latency requirements. Training large RNN-based MTPPs under these constraints was unsuccessful in practice—we faced frequent restarts or truncated backpropagation windows, reducing the model's efficacy.

5 Discussion and (Partial) Lessons Learned

Tensor Decomposition (Partial Fix). A subsequent attempt used tensor decomposition in the RNN's weight matrices to control parameter blow-up (Oseledets, 2011; Novikov et al., 2015). This improved memory usage and partially alleviated overfitting, but introduced its own tuning complexities (e.g., choosing the TT-rank).

Needs Domain-Driven Architecture. We suspect that domain constraints, such as feature grouping (e.g., treating each instrument's features together), or physically meaningful cross-terms, would regularize the model more effectively than purely black-box RNNs.

Future Considerations. Even with compression, training remains sensitive to hyperparameters and partial distribution shifts. Real-time online updates and robust confidence intervals are still open challenges for neural MTPPs in finance.

6 Conclusion

We presented a real-world deployment attempt of RNN-based MTPPs in high-dimensional finance and discovered issues of parameter inflation, data mismatch, and performance instability that are not highlighted in typical benchmarks. Our partial fix with tensor decomposition points to the broader need for *domain-driven neural architectures* and specialized regularization when applying deep learning solutions to complex real-world event streams.

Acknowledgments

We thank colleagues in quantitative research and ML engineering for helpful discussions.

References

Nan Du, Hanjun Dai, Rakshit Trivedi, Utkarsh Upadhyay, Manuel Gomez-Rodriguez, and Le Song. Recurrent marked temporal point processes: embedding event history to vector. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1555–1564, 2016.

Jian Mei and Jason Eisner. Neural hawkes process: A neurally self-modulating multivariate point process. In *Advances in Neural Information Processing Systems (NIPS)*, 2017.

Alexander Novikov, Dmitry Podoprikhin, Anton Osokin, and Dmitry Vetrov. Tensorizing neural networks. In *Advances in Neural Information Processing Systems (NIPS)*, pp. 442–450, 2015.

Ivan V. Oseledets. Tensor-train decomposition. *SIAM Journal on Scientific Computing*, 33(5):2295–2317, 2011.