A SYSTEMATIC LITERATURE REVIEW ON MACHINE LEARNING AND DEEP LEARNING FOR REAL TIME TRADING IN FINANCIAL MARKETS

University of South Wales
Prifysgol De Cymru

By Pradeep Kumar Reddy Asi

Student ID: 30118327

1. Abstract

This project is based on the planning, design, implementation and evaluation of machine learning and deep learning models for real-time algorithmic trading in financial domain. The initial objective is to build and establish intelligent prediction models and algorithms capable of analysing historical market data to find patterns, predict market prices and support automated trading decisions in high frequency trading. To improve prediction performance of the model's technical indicators like RSI, MACD, ATR, ADX, Bollinger Bands are implemented which indicate momentum, volatility and trend strength. These are computed as informative features. Many types of market trend and direction of future price are implemented through machine learning models like Random Forest, SVM (Support Vector Machine), Logistic Regression, XGBoost, and LightGBM etc. Deep learning models like LSTM, CNN, and GRU are employed to capture the temporal correlation in price sequences and generate accurate forecasts. The performance of these models is continuously tested and evaluated by back testing on historical data, replicating real-world trading scenarios. Crucial financial metrics like return on investment, Sharpe ratio, and maximum drawdown are incorporated into this project to assess adjustment of returns based on risks, stability, and profitability.

2. INTRODUCTION

In this project, the literature review checks the current state of research is machine learning and deep learning for real time algorithmic trading. Financial markets are sometimes complex and data-driven, there is growing demand for smart systems which can predict trends, optimize strategies and execute trades. This review is based on ML and DL based autonomous agents, financial data preprocessing, model design, model evaluation and ethical considerations.

3. MACHINE LEARNING IN MARKET PREDICTION

Machine Learning algorithms are broadly used for stock market prediction addressing to uncover patterns in historical price movements. Classical statistical models sometimes fall short in assimilating the complex, nonlinear and changing behaviour of financial markets. Machine Learning methods like Support Vector Machines (SVM), Random Forest (RF) and Long Short-Term Memory Networks (LSTM) have shown promising capabilities in studying intricate patterns and adjusting to new data.

Support Vector Machines (SVM) and Random Forest are considered some of the best supervised models, because of their resilience and capacity for generalization. Research shows that ensemble techniques like RF outperform linear models when it comes to prediction accuracy, notably when technical indicators are used as input features. SVM has shown efficient in handling high-dimensional and noisy data.

Random Forest and SVM models require effective feature engineering, because their performance tends to decrease when faced with highly volatile market data.

Current advancement in deep learning models like LSTM, are built to capture temporal dependencies in sequential data. These approaches have outperformed conventional classifiers on benchmark datasets by efficiently modelling the autocorrelation exists in financial data. As these models have high predictive power, yet these models need large datasets and are prone to overfitting, when market conditions change rapidly.

Rank	Model/Algorithm	Applications	Strength	Limitations
1	LightGBM/XGBoost	Signal prediction, feature selection	Strong accuracy, Handles missing data	Needs nice features
2	Random Forest	Risk Modelling, Direction Prediction	Robust, interpretable perform well with little tuning	Less accurate than boosting
3	SVM	Mean Reversion Strategies, Trend Classification	Nice for clean datasets, margin maximization	Poor scalability
4	Logistic Regression	Binary Outcome (Up/Down) prices	Simple and fast	Weak with nonlinear patterns
5	Naïve Bayes	Regime switching, event detection	Quick to implement	Not ideal for noisy data
6	Decision Trees	Alpha modelling	Good with categorical and numerical values	Slower than LightGBM/XGBoost
7	Ensemble Methods	Diversified signal models	Improves stability	Complexity in implementation

Table 1: Ranking of Machine Learning Models in Financial Trading

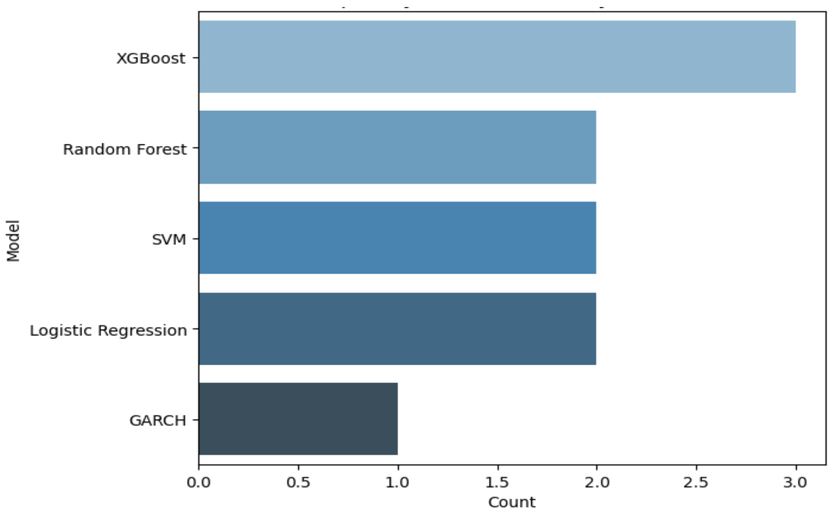

Figure 1: Frequency of ML models in Financial Trading

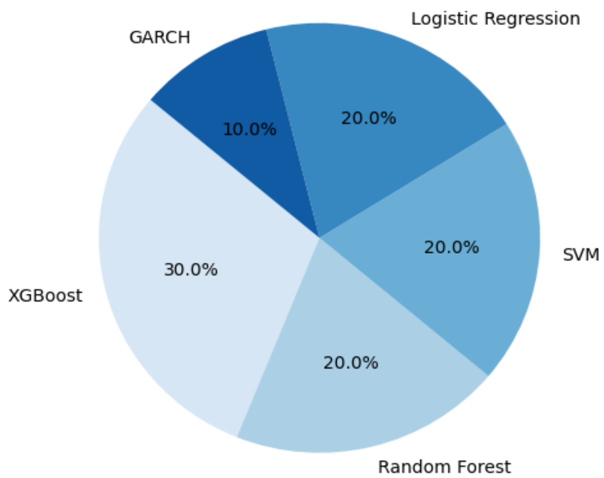

Figure 2: Pie Chart showing the usage of ML models in Financial Trading

4. DEEP LEARNING METHODS IN TRADING

Deep Learning appeared as a strong method in algorithmic trading by enabling autonomous agents to study optimal trading strategies through interaction with real market data. Unlike conventional methods, deep learning methods do not require labelled data but instead depend on a reward function that evaluates actions based on financial performance metrics. This regulates well with the aim of maximizing long-term in financial markets.

Many Deep Learning algorithms are applied to trading tasks. For example, Deep Q-Networks (DQN), have concluded effectiveness in discrete-action trading conditions. The Deep Q-Networks combine Q-learning with deep neural networks. Proximal Policy

Optimization (PPO) and Actor-Critic methods like A2C and A3C are better for continuous action spaces and is best for stable trading strategies.

The Deep Learning methods are capable in complex and dynamic market scenarios. Despite its ability there are some risks involved as well. Deep Learning agents requires careful tuning of reward functions and hyperparameters to avoid volatility to suboptimal policies. Real world deployment is strained by the reality gap between simulated and live trading scenarios. Studies shows high profitability in back tests, but these outcomes sometimes fail to generalize due to overfitting to historical data and shortage of consideration for latency, slippage and transaction costs.

Some deep learning models integrate Sharpe ratio into the reward structure, many still revise raw returns, which can result in volatile strategies. Most Deep Learning agents operate as black boxes, which makes difficult for stakeholders to audit their decision-making processes.

To overcome these issues, the use of PPO for continuous trading is useful to integrate risk related adjustment of performance metrics into the training reward function.

Rank	Model/Algorithm	Application	Strength	Limitations
1	Transformer (BERT, FinBERT)	Sentiment Analysis, order book modelling, forecasting	Best for long range dependencies, multi-modal data	Resource intensive, complex tuning
2	LSTM (Long Short-term Memory)	Time Series Forecasting, Volatility Prediction	Nice for sequential financial data	Slow training, noisy for long sequences
3	GRU (Gated Recurrent Unit)	Alternative to LSTM, Trend Modelling	Nice temporal memory	Less expensive than LSTM in some cases
4	Deep Reinforcement Learning (DQN, PPO, A3C)	Trade Execution, Portfolio Optimization	Learns using environment interaction	Needs careful reward engineering
5	CNN(1D/2D/3D)	Price pattern recognition	Best at feature extraction from raw LOBs	Limited memory of past events
6	Autoencoders	Dimensionality Reduction, anomaly Detection	Useful for feature compression	Poor in predictive performance.

Table 2: Top Deep Learning models in financial trading

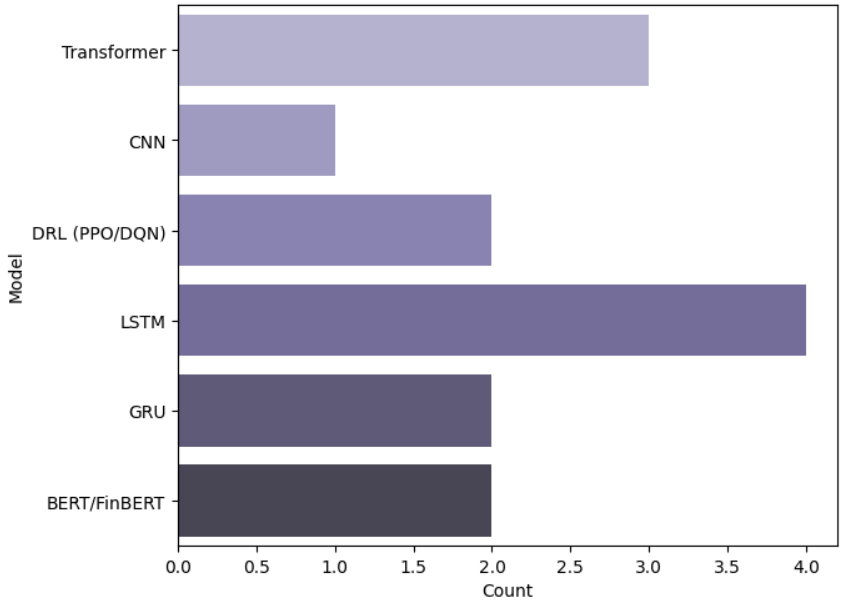

Figure 3: Frequency of DL models in Financial Trading

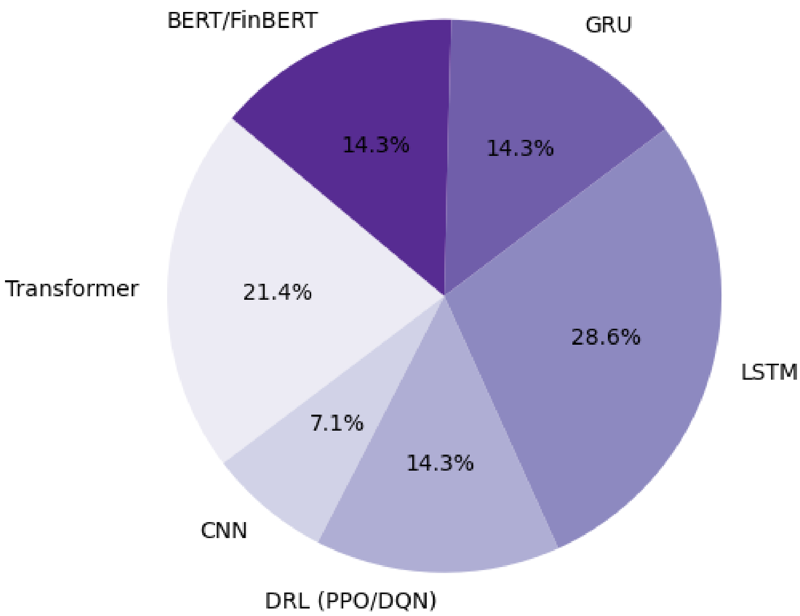

Figure 4: Pie Chart showing the usage of DL models in Financial Trading

5. FINANCIAL DATA PREPROCESSING

Efficient and reliable data preprocessing is a crucial as well as critical component of algorithmic trading, as financial time series are sometimes noisy and incomplete. Preprocessing usually involves handling missing values, normalization and feature engineering to build descriptive inputs for machine learning and deep learning models.

Time-series forecasting needs the use of techniques like log transformations, differencing and rolling statistics to stabilize variance. Additionally, resampling techniques are often deployed to coordinate with multiple data streams like indicators, sentiment metrics and prices.

Feature engineering is especially impactful in financial markets. Mostly used features include RSI, MACD, ATR, moving averages, volatility indicators and sentiment scores derived from financial news. For sentiment analysis models like FinBERT is useful to extract polarity signals from text data, increasing model performance on major market events.

Financial markets are highly volatile and dynamic and features that may be predictable in one regime may become obsolete in other. Recent approaches like automated feature selection or online learning methods is adaptable to changing conditions.

6. EVALUATION METRICS AND BACKTESTING

Evaluation of the performance of trading algorithms and decisions involves more than just predictive accuracy. It demands both risk and profitability. Financial metrics like Sortino ratio, Sharpe ratio, maximum drawdown and cumulative return are useful for determining the real-world application on trading market data.

Back testing which shows how a strategy would perform on historical data is most common method for evaluation. It gives insights on profitability, risk, robustness and exposure. Poorly built back test may lead to misleading results because of issues such as overfitting and impractical belief about transaction costs and liquidity.

To overcome these risks researchers, prefer the use of cross-validation over rolling windows, walk-forward analysis and establishment of realistic trading constraints. Including risk-adjusted returns in model objectives while training can help adjust learning with actual trading goals.

In this project, robust back testing frameworks ensure that ML and DL models will not only perform well on historical data but also generalize to future market conditions. These evaluations will have realistic assumptions about market latency, commissions and slippage.

7. ETHICAL, LEGAL AND REGULATORY CONSIDERATIONS

The increase in adoption of artificial intelligence techniques in financial markets raises many ethical, standard and regulatory challenges. General Data Protection Regulation (GDPR) is essential when handling sensitive information. In addition, AI or software developers must focus on issues of informed consent, when using public datasets, to make sure ethical standards are met.

Another concern is the possibility of using AI related strategies to contribute to market manipulation. For instance, high frequency trading algorithms earlier linked to sudden market downturns. Such continuous testing and ethical guidelines are needed to avoid unintended consequences.

The research will follow best rules and practices, which includes explanation of algorithmic trading, risk controls and compliance with related financial regulations. This research considers broader impact of establishing autonomous trading agents and involve safeguards to prevent instability in market behaviour.

8. Critical Gaps and Project Contribution

Existing research explains the potential of ML and DL algorithmic trading, but many critical gaps persist. Firstly, few numbers of studies focus on predictive accuracy without deploying risk-adjusted performance metrics, which are crucial for practical trading. This shows that models that are profitable in one theory but may me volatile to another.

Secondly most DRL implementations have certain limitations to simulated environments with simplified hypothesis and hardly transition into live trading. The gap between historical back tests and real time trading gives a challenge in terms of stability, robustness, latency management and generalization. Additionally, many of this research give extra focus on practical aspects like transaction costs, slippage and real scenario conditions like liquidity limits and order execution delays.

Another issue is the lack of focus on market regime shifts. Financial markets are constantly moving and evolving, and static models may not fit or may become obsolete. Very limited studies explore learning or model retraining works which adapt to evolving situations. Furthermore, explainability and openness are addressed to minimum, making these systems little trustworthy to business stakeholders and regulators.

The project aims to fill these gaps by building a hybrid ML-DL framework optimized for real time algorithmic trading. The models will be examined using predictive and financial metrics, and training scenarios will be enhanced with realistic constraints. By using these methods system will incorporate basic explainability modules to increase trustworthiness and interpretability.

9. Conclusion

In this literature review machine learning and deep learning methods have been compiled in algorithmic trading. It has examined and evaluated the strengths and weaknesses of conventional methods for forecasting the market. The use of deep learning methods in the creation of autonomous trading agents, and the necessity of reliable data preparation methods. It also highlighted how important back testing techniques and financial performance indicators for assessing trading strategies. It also pointed out ethical and regulatory factors such as data compliance and algorithm transparency for careful and responsible AI implementation in the financial sector.

By addressing these limitations, this research will provide a practical solution to real-time trading using AI. The proposed model will combine robust model architectures with realistic evaluation, which results to deliver a reliable and high-performance trading system.

10. References

[1] Lin, C.Y. and Lobo Marques, J.A. (2024). Stock market prediction using artificial intelligence: A systematic review of systematic reviews. *Social Sciences & Humanities Open*, [online] 9, p.100864. doi:https://doi.org/10.1016/j.ssaho.2024.100864.

[2] Millea, A. (2021). Deep Reinforcement Learning For Trading—A Critical Survey. *Data*, 6(11), p.119. doi:https://doi.org/10.3390/data6110119.

[3] Yu, Y. (2024). A Survey of Deep Reinforcement Learning in Financial Markets. *Atlantis Highlights in Computer Sciences/Atlantis highlights in computer sciences*, [online] pp.188–194. doi:https://doi.org/10.2991/978-94-6463-419-8_24.

[4] Mohammadshafie, A., Mirzaeinia, A., Jumakhan, H. and Mirzaeinia, A. (n.d.). *Deep Reinforcement Learning Strategies in Finance: Insights into Asset Holding, Trading Behavior, and Purchase Diversity Regular Research Paper (CSCE-ICAI'24)*.

[5] Zeng, Z., Kaur, R., Siddagangappa, S., Rahimi, S., Balch, T., Veloso, M. and Morgan, J. (n.d.). *Financial Time Series Forecasting using CNN and Transformer*.

[6] Bilokon, P. and Qiu, Y. (2023). *TRANSFORMERS VERSUS LSTMS FOR ELECTRONIC TRADING A PREPRINT Transformers versus LSTMs for electronic trading A PREPRINT*.

[7] Buczyński, M., Chlebus, M., Kopczewska, K. and Zajenkowski, M. (2023). Financial Time Series Models—Comprehensive Review of Deep Learning Approaches and Practical Recommendations. *ITISE 2023*, p.79. doi:https://doi.org/10.3390/engproc2023039079.

[8] Murat Ozbayoglu, A., Gudelek, M. and Sezer, O. (n.d.). *Deep Learning for Financial Applications : A Survey*.

[9] Xie, L., Chen, Z. and Yu, S. (2024). Deep Convolutional Transformer Network for Stock Movement Prediction. *Electronics*, 13(21), p.4225. doi:https://doi.org/10.3390/electronics13214225.

[10] Olorunnimbe, K. and Viktor, H. (2022). Deep learning in the stock market—a systematic survey of practice, backtesting, and applications. *Artificial Intelligence Review*, 56(3), pp.2057–2109. doi:https://doi.org/10.1007/s10462-022-10226-0.

[11] Kabir, M.R., Bhadra, D., Ridoy, M. and Milanova, M. (2025). LSTM–Transformer-Based Robust Hybrid Deep Learning Model for Financial Time Series Forecasting. *Sci*, 7(1), p.7. doi:https://doi.org/10.3390/sci7010007.

[12] Ganesh, P. and Rakheja, P. (n.d.). *VLSTM: VERY LONG SHORT-TERM MEMORY NETWORKS FOR HIGH-FREQUENCY TRADING*.

[13] Barez, F., Bilokon, P., Gervais, A. and Lisitsyn, N. (2023). Exploring the Advantages of Transformers for High-Frequency Trading. *SSRN Electronic Journal*. doi:https://doi.org/10.2139/ssrn.4364833.

[14] Tran, D.T., Iosifidis, A., Kanniainen, J. and Gabbouj, M. (2019). Temporal Attention-Augmented Bilinear Network for Financial Time-Series Data Analysis. *IEEE Transactions on Neural Networks and Learning Systems*, 30(5), pp.1407–1418. doi:https://doi.org/10.1109/tnnls.2018.2869225.

[15] Zhang, C., Sjarif, N.N.A. and Ibrahim, R. (2023). Deep learning models for price forecasting of financial time series: A review of recent advancements: 2020–2022. *WIREs Data Mining and Knowledge Discovery*, 14(1). doi:https://doi.org/10.1002/widm.1519.

[16] Mienye, E., Jere, N., Obaido, G., Mienye, I.D. and Aruleba, K. (2024). Deep Learning in Finance: A Survey of Applications and Techniques. *AI*, 5(4), pp.2066–2091. doi:https://doi.org/10.3390/ai5040101.

[17] Gao, H., Kou, G., Liang, H., Zhang, H., Chao, X., Li, C.-C. and Dong, Y. (2024). Machine learning in business and finance: a literature review and research opportunities. *Financial Innovation*, 10(1). doi:https://doi.org/10.1186/s40854-024-00629-z.

[18] Subasi, A., Amir, F., Bagedo, K., Shams, A. and Sarirete, A. (2021). Stock Market Prediction Using Machine Learning. *Procedia Computer Science*, 194, pp.173–179. doi:https://doi.org/10.1016/j.procs.2021.10.071.

[19] Wu, R. (2024). Leveraging Deep Learning Techniques in High-Frequency Trading: Computational Opportunities and Mathematical Challenges. *Published By SOUTHERN UNITED ACADEMY OF SCIENCES*, 2(4). doi:https://doi.org/10.5281/zenodo.12747424.

[20] Borkar, S.V. and Sangve, S.M. (2025). A Critical Analysis on Anomaly Detection in High-Frequency Financial Data Using Deep Learning for Options. doi:https://doi.org/10.20944/preprints202505.0199.v1.

[21] Briola, A., Turiel, J., Marcaccioli, R., Cauderan, A. and Aste, T. (n.d.). *Deep Reinforcement Learning for Active High Frequency Trading*.

[22] Kumar, P., Khan, E. and Gönen, M. (n.d.). Deep Reinforcement Learning for High-Frequency Market Making. *Proceedings of Machine Learning Research*, 189, pp.2022–2022.

[23] Zong, C., Wang, C., Qin, M., Feng, L., Wang, X. and An, B. (2024). MacroHFT: Memory Augmented Context-aware Reinforcement Learning On High Frequency Trading. *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp.4712–4721. doi:https://doi.org/10.1145/3637528.3672064.

[24] Qin, M., Sun, S., Zhang, W., Xia, H., Wang, X. and An, B. (2024). EarnHFT: Efficient Hierarchical Reinforcement Learning for High Frequency Trading. *Proceedings of the AAAI Conference on Artificial Intelligence*, 38(13), pp.14669–14676. doi:https://doi.org/10.1609/aaai.v38i13.29384.

[25] Guilbaud, F. and Pham, H. (2011). Optimal High Frequency Trading with Limit and Market Orders. *SSRN Electronic Journal*. doi:https://doi.org/10.2139/ssrn.1871969.

[26] Petukhina, A.A., Reule, R.C.G. and Härdle, W.K. (2020). Rise of the machines? Intraday high-frequency trading patterns of cryptocurrencies. *The European Journal of Finance*, 27(1-2), pp.8–30. doi:https://doi.org/10.1080/1351847x.2020.1789684.

[27] Zhao, Z., Zhang, X., Wen, J., Liu, M. and Ma, X. (n.d.). *Label Unbalance in High-frequency Trading*.

[28] Bilokon, P. and Gunduz, B. (2023). C Design Patterns for Low-Latency Applications Including High-Frequency Trading. *SSRN Electronic Journal*. doi:https://doi.org/10.2139/ssrn.4565813.

[29] Kisiel, D. and Gorse, D. (n.d.). *Axial-LOB: High-Frequency Trading with Axial Attention*.

[30] Zhou, L., Qin, K., Ferreira Torres, C., Le, D. and Gervais, A. (n.d.). *High-Frequency Trading on Decentralized On-Chain Exchanges*.

[31] Rahman, A. and Upadhye, N. (2024). *HYBRID VECTOR AUTO REGRESSION AND NEURAL NETWORK MODEL FOR ORDER FLOW IMBALANCE PREDICTION IN HIGH-FREQUENCY TRADING*.

[32] Murphy, N.J. and Gebbie, T.J. (2021). Learning the dynamics of technical trading strategies. *Quantitative Finance*, 21(8), pp.1325–1349. doi:https://doi.org/10.1080/14697688.2020.1869292.

[33] Taghian, M., Asadi, A. and Safabakhsh, R. (n.d.). *A Reinforcement Learning Based Encoder-Decoder Framework for Learning Stock Trading Rules*.

[34] Berti, L., Prenkaj, B. and Velardi, P. (n.d.). *TRADES: Generating Realistic Market Simulations with Diffusion Models*.

[35] Ibikunle, G., Moews, B., Muravyev, D. and Rzayev, K. (n.d.). *Data-Driven Measures of High-Frequency Trading*.

[36] Yagi, I., Masuda, Y. and Mizuta, T. (2020). Analysis of the Impact of High-Frequency Trading on Artificial Market Liquidity. *IEEE Transactions on Computational Social Systems*, 7(6), pp.1324–1334. doi:https://doi.org/10.1109/tcss.2020.3019352.

[37] Lim, Y.-S. and Gorse, D. (n.d.). *Deep Probabilistic Modelling of Price Movements for High-Frequency Trading*.

[38] Leal, L., Lauriere, M. and Lehalle, C.-A. . (2022). Learning a functional control for high-frequency finance. *Quantitative Finance*, 22(11), pp.1973–1987. doi:https://doi.org/10.1080/14697688.2022.2106885.

[39] Goudarzi, M. and Bazzana, F. (2023). Identification of high-frequency trading: A machine learning approach. *Research in International Business and Finance*, 66, p.102078. doi:https://doi.org/10.1016/j.ribaf.2023.102078.

[40] Ntakaris, A. and Ibikunle, G. (n.d.). *Minimal Batch Adaptive Learning Policy Engine for Real-Time Mid-Price Forecasting in High-Frequency Trading*.

[41] Li, C., Shen, L. and Qian, G. (2023). Online Hybrid Neural Network for Stock Price Prediction: A Case Study of High-Frequency Stock Trading in the Chinese Market. *Econometrics*, 11(2), p.13. doi:https://doi.org/10.3390/econometrics11020013.

[42] Bao, Q., Wang, J., Gong, H., Zhang, Y. and Feng, H. (n.d.). *A Deep Learning Approach to Anomaly Detection in High-Frequency Trading Data*.

[43] Jaddu, K. and Bilokon, P. (2023). Combining Deep Learning on Order Books with Reinforcement Learning for Profitable Trading. *SSRN Electronic Journal*. doi:https://doi.org/10.2139/ssrn.4611708.

[44] Sarkar, S. (2023). *Harnessing Deep Q-Learning for Enhanced Statistical Arbitrage in High-Frequency Trading: A Comprehensive Exploration*.

[45] Bhatia, S., Peri, S., Friedman, S. and Malen, M. (2024). *High-Frequency Trading Liquidity Analysis Application of Machine Learning Classification*.

[46] Yang, Y. (n.d.). Deep Learning-Driven Order Execution Strategies in High-Frequency Trading: An Empirical Study on Enhancing Market Efficiency. doi:https://doi.org/10.54254/2755-2721/118/2025.18469.

[47] Nagy, P., Calliess, J.-P. and Zohren, S. (2023). Asynchronous Deep Double Dueling Q-learning for trading-signal execution in limit order book markets. *Frontiers in Artificial Intelligence*, 6. doi:https://doi.org/10.3389/frai.2023.1151003.

[48] Liu, F. and Tian, Y. (2024). Deep Reinforcement Learning-based Algorithmic Optimisation and Risk Management for High Frequency Trading. *Journal of Computing and Electronic Information Management*, 14(1), pp.28–32. doi:https://doi.org/10.54097/fgblm2ei.

[49] Kumar, P. (2024). Deep Hawkes process for high-frequency market making. *Journal of Banking and Financial Technology*, 8(1), pp.11–28. doi:https://doi.org/10.1007/s42786-024-00049-8.

[50] Hou, Y. (2024). Predictive modeling in high-frequency trading using machine learning. Applied and Computational Engineering, 90(1), pp.61–65. doi:https://doi.org/10.54254/2755-2721/90/20241764.

[51] Chen, Y., Li, M., Shu, M., Bi, W. and Xia, S. (2024). Multi-modal Market Manipulation Detection in High-Frequency Trading Using Graph Neural Networks. *Journal of Industrial Engineering and Applied Science*, 2(6), pp.111–120. doi:https://doi.org/10.70393/6a69656173.323432.

[52] Li, M., Shu, M. and Lu, T. (2024). Anomaly Pattern Detection in High-Frequency Trading Using Graph Neural Networks. *Journal of Industrial Engineering and Applied Science*, 2(6), pp.77–85. doi:https://doi.org/10.70393/6a69656173.323430.