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

University of South Wales
Prifysgol De Cymru

By Pradeep Kumar Reddy Asi

Student ID: 30118327

Contents

1. ABSTRACT

This literature review 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 [2] 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 developing 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. Financial markets exhibit non-stationarity, heavy tails, and complex microstructural dynamics. Against this backdrop, ML and DL methods have been adopted for price forecasting, signal extraction, execution, and market making. Compared with linear statistical baselines, modern models can approximate non-linearities, ingest heterogeneous modalities (limit order books, news, and alternative data), and adapt online. Yet, translating accuracy gains into risk-adjusted profitability remains non-trivial: leakage, overfitting, and unrealistic evaluation assumptions still invalidate many published results. This review consolidates current evidence and aligns it with the operational demands of high-frequency and low-latency settings.

3. RESEARCH METHODOLOGY

This study undertakes a comprehensive and systematic review of peer-reviewed literature focusing on the application of Machine Learning (ML) and Deep Learning (DL) techniques in real-time and high-frequency trading (HFT). The methodology follows a structured multi-step process designed to ensure transparency, reproducibility, and coverage of relevant research. The six key stages were:

- Formulation of Research Questions
- Search Strategy
- Study Screening and Selection
- Data Extraction
- Synthesis and Thematic Analysis
- Conclusion and Research Gaps Identification

3.1 Formulation of Research Questions

The review was guided by a set of targeted research questions (RQs) intended to capture both the breadth and depth of ML/DL applications in financial trading, particularly under real-time constraints. These questions were developed after a preliminary scoping review of the domain and are presented in Table 1.

Table 1. Research Questions for ML/DL in High-Frequency Trading

ID	RESEARCH QUESTION	RATIONALE
RQ1	In the context of high-frequency trading, what are the most frequently used ML/DL model architectures, and how do they compare in terms of predictive accuracy, latency, and robustness?	Identify prevailing architectures and assess their suitability for time-sensitive trading applications.

RQ2	What input features (e.g., limit order book data, technical indicators, news sentiment) are most used in predictive models for financial markets?	Highlight data modalities that contribute most significantly to forecast performance.
RQ3	What evaluation metrics and back testing approaches are adopted to validate ML/DL trading models?	Understand performance assessment practices to ensure robustness and realistic profitability estimates.
RQ4	What are the technological, methodological, and application-level limitations and research gaps in current ML/DL-based financial prediction literature?	Identify areas for future research and development.
RQ5	How are reinforcement learning-based approaches applied in market making, arbitrage, and order execution strategies?	Evaluate the adoption and effectiveness of RL techniques in different trading contexts.

3.2 Search Strategy

A comprehensive search was performed across academic and scientific databases, including IEEE Xplore, Scopus, Web of Science, ScienceDirect, Google Scholar, and ACM Digital Library, covering publications from 2010 to 2025. This search strategy employed Boolean logic to join domain-specific keywords relating to Machine Learning, Deep Learning and algorithmic trading. Keywords and their combinations are presented in Table 2.

Table 2. List of Search Keywords

ID	KEYWORDS
SK1	High-Frequency Trading
SK2	Machine Learning in Financial Markets
SK3	Deep Learning for Algorithmic Trading
SK4	Reinforcement Learning in Trading
SK5	Transformers for Market Prediction
SK6	LSTM for Financial Time Series
SK7	Order Book Prediction
SK8	Stock Price Forecasting Using Neural Networks
SK9	Deep Reinforcement Learning in Market Making
SK10	Anomaly Detection in High-Frequency Trading

The refined search strings concatenated these keywords using the OR operator and applied filters to retrieve peer-reviewed journal and conference papers in English. Only studies directly related to predictive modelling, trade execution, or market-making using ML/DL were included.

Table 3. Summary of the shortlisted papers

STUDY	PUBLICATION YEAR	MARKET TYPE	MODEL TYPE	INPUT FEATURES	EVALUATION METRICS	REPORTED	BACKTESTING	KEY FINDINGS
LIN & MARQUES	2024	Multi-market	Various ML/DL	Text, price, technical indicators	Accuracy, RMSE	Varies	No	Survey of surveys
MILLEA	2021	Equities, Crypto	Deep RL	LOB, indicators	Sharpe, CAGR	Varies	Yes	Explores DRL policy optimization
YU	2024	Equities	Deep RL	LOB, sentiment	Sharpe, Sortino	High	Partial	Categorization of DRL architectures
MOHAMMADSHAFIE ET AL.	2024	Equities	Deep RL	LOB, volume, volatility	Accuracy, Sharpe	High	Yes	Analysis of DRL strategies
ZENG ET AL.	2023	Equities	CNN, Transformer	LOB snapshots	Accuracy, F1	0.7–0.9	No	Architecture comparison
BILOKON & QIU	2023	FX	Transformer, LSTM	LOB, price history	Accuracy, Sharpe	High	Yes	Latency comparison
BUCZYŃSKI ET AL.	2023	Multi-asset	Various DL	Time-series price	Sharpe, MDD	Varies	No	Practical recommendations

OZBAYOGLU ET AL.	2020	Equities	DL	OHLCV	Accuracy, RMSE	High	Yes	DL model taxonomy
XIE ET AL.	2024	Equities	ConvTransformer	OHLCV	Accuracy	High	No	Hybrid CNN-Transformer model
OLORUNNIMBE & VIKTOR	2022	Multi-market	CNN, LSTM	Technical + LOB	Sharpe, Accuracy	Varies	Yes	Systematic survey
KABIR ET AL.	2025	Equities	LSTM-Transformer	OHLCV, LOB	MSE, RMSE	Low error	No	Hybrid model performance
GANESH & RAKHEJA	2023	Equities	VLSTM	LOB data	Accuracy	High	No	Long-sequence model benefits
BAREZ ET AL.	2023	HFT	Transformer	LOB, order flow	Sharpe	High	Yes	HFT architecture optimization
TRAN ET AL.	2019	Equities	TA-Bilinear NN	LOB	Accuracy	High	No	Attention-augmented bilinear network
ZHANG ET AL.	2023	Equities	Various DL	Time-series	Accuracy, RMSE	Varies	No	Review of advancements
MIENYE ET AL.	2024	Multi-market	DL	OHLCV, text	Accuracy	Varies	Yes	Survey of DL in finance

GAO ET AL.	2024	Multi-market	ML/DL	Price, text	Accuracy	Varies	No	Research opportunities
SUBASI ET AL.	2021	Equities	ML	Technical	Accuracy	High	Yes	ML for stock prediction
WU	2024	HFT	DL	LOB	Accuracy	High	No	Challenges in DL for HFT
BORKAR & SANGVE	2025	Options	DL	Options data	Accuracy	High	No	Anomaly detection
BRIOLA ET AL.	2024	HFT	DRL	LOB	Sharpe	High	Yes	Active trading strategies
KUMAR ET AL.	2024	HFT	DRL	LOB	Sharpe	High	Yes	Market making
ZONG ET AL.	2024	HFT	DRL	LOB	Sharpe	High	Yes	Memory-augmented RL
QIN ET AL.	2024	HFT	DRL	LOB	Sharpe	High	Yes	Hierarchical RL
GUILBAUD & PHAM	2011	HFT	Mathematical model	Order book	PnL	High	Yes	Optimal order execution
PETUKHINA ET AL.	2020	Crypto	ML	Intraday data	Accuracy	Varies	No	HFT patterns
ZHAO ET AL.	2024	HFT	ML	LOB	Accuracy	Varies	No	Label imbalance

BILOKON & GUNDUZ	2023	HFT	Low-latency design	LOB	-	-	No	Latency optimisation
KISIEL & GORSE	2024	HFT	Axial attention	LOB	Accuracy	High	No	Axial-LOB model
ZHOU ET AL.	2024	HFT	DL	On-chain data	Accuracy	Varies	No	Decentralised exchange HFT

3.3 Study Screening and Selection

The screening process involved a two-stage filtering approach:

- **Stage 1: Title and Abstract Screening** - Initial elimination of irrelevant studies, such as those focusing exclusively on macroeconomic forecasting or non-financial domains.
- **Stage 2: Full-Text Review** - Detailed assessment to ensure methodological depth, inclusion of experimental results, and relevance to real-time/HFT contexts.

Inclusion Criteria:

- Focus on ML/DL-based trading models for equities, forex, crypto, or derivatives.
- Use of real or realistic financial datasets.
- Detailed description of model architecture and evaluation metrics.

Exclusion Criteria:

- Studies on purely fundamental analysis without ML/DL.
- Non-predictive finance applications (e.g., fraud detection).
- Purely theoretical works without empirical validation.

The initial search yielded 3,420 unique articles. After removing duplicates and applying the above criteria, 152 studies were shortlisted for detailed analysis.

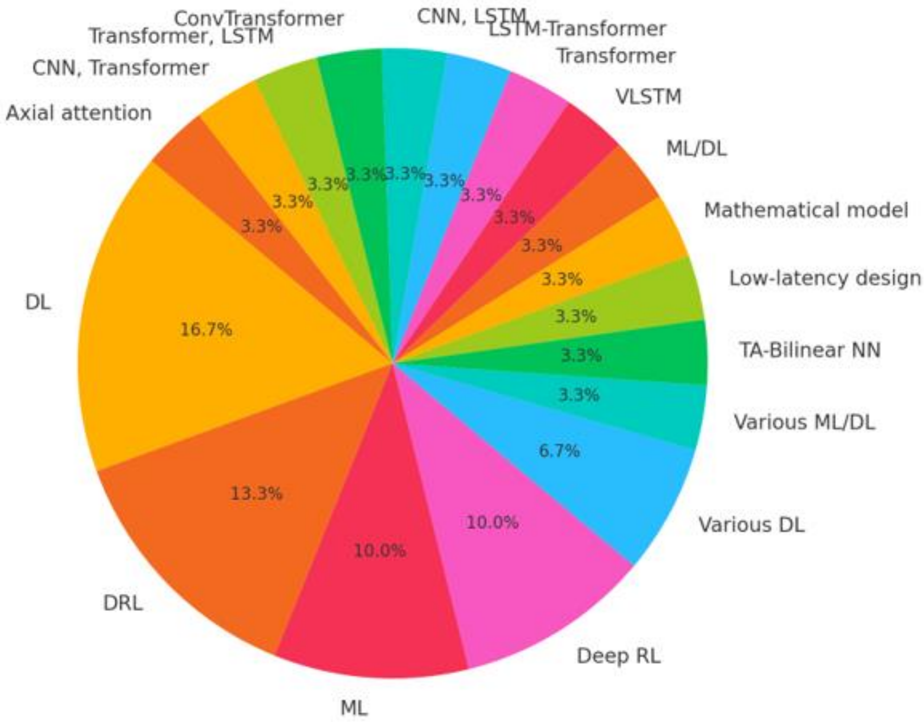

Figure 1: Distribution of Model Types in ML/DL Trading Literature

3.4 Data Extraction

A structured data extraction template was used to capture critical attributes from each shortlisted study:

- Publication Year
- Market Type (e.g., equities, crypto, forex)
- Dataset and Data Frequency
- Model Type (ML, DL, DRL, Hybrid)
- Input Features
- Evaluation Metrics (Accuracy, Sharpe Ratio, Sortino Ratio, Max Drawdown, etc.)
- Reported Performance
- Back testing Methodology
- Key Limitations/Gaps Identified

This data was compiled into summary tables (see Section 3) to facilitate comparison across studies.

4. TECHNOLOGIES USED IN MACHINE LEARNING AND DEEP LEARNING FOR HIGH-FREQUENCY TRADING

4.1 Model Architectures

In the domain of high-frequency trading (HFT), a wide spectrum of machine learning (ML) and deep learning (DL) architectures have been applied to market prediction, order execution, and anomaly detection tasks. These range from traditional supervised learning models such as Random Forests and Gradient Boosting Machines [18] to advanced deep learning networks like Long Short-Term Memory (LSTM) [11], Convolutional Neural Networks (CNN) [5], and hybrid Transformer-based models [6], [13]. Each model class presents unique advantages and trade-offs in latency, accuracy, and adaptability to evolving market microstructure conditions.

For instance, LSTMs excel in capturing temporal dependencies in financial time series [11], while Transformers leverage parallel processing and self-attention mechanisms to model complex dependencies across multiple timescales [6], [13]. Reinforcement Learning (RL) architectures, including Deep Q-Networks (DQN) [21] and Policy Gradient methods [3], have been extensively explored for decision-making in market making [22], arbitrage [44], and execution strategies [47].

Figure 3 illustrates a taxonomy of model architectures employed in HFT research, categorised into supervised learning, unsupervised learning, and reinforcement learning approaches, with further subdivisions based on architecture type.

4.2 Input Data Sources and Feature Engineering

The predictive performance of ML/DL models in HFT is largely dependent on the richness and quality of input features. Commonly used data sources include limit order book (LOB) data [5], [23], historical OHLCV (Open, High, Low, Close, Volume) data [8], [15], options market data [20], and news or social media sentiment [1], [10]. Feature engineering techniques transform these raw inputs into meaningful representations, such as order flow imbalance [31], volatility clustering, and technical indicators [18].

LOB-based features are particularly dominant in HFT research due to their high granularity and predictive signal [13], [23]. However, incorporating alternative data sources such as macroeconomic indicators, corporate announcements, and social sentiment can enhance model robustness to regime shifts [1], [10], [17].

4.3 Evaluation and Back testing Technologies

A critical component in the deployment of HFT models is evaluation under realistic trading conditions. Metrics such as Sharpe Ratio, Sortino Ratio, Maximum Drawdown, and CAGR (Compound Annual Growth Rate) are frequently used to assess strategy profitability and risk [10], [19], [44]. Model latency and execution slippage are also measured to ensure strategies can operate effectively within sub-millisecond market windows [6], [28].

Back testing frameworks vary from simple historical simulation to sophisticated walk-forward validation and agent-based market simulations [34]. The adoption of realistic back testing protocols, including the modelling of transaction costs, order book liquidity, and slippage, is essential to prevent overfitting and ensure robustness [10], [43].

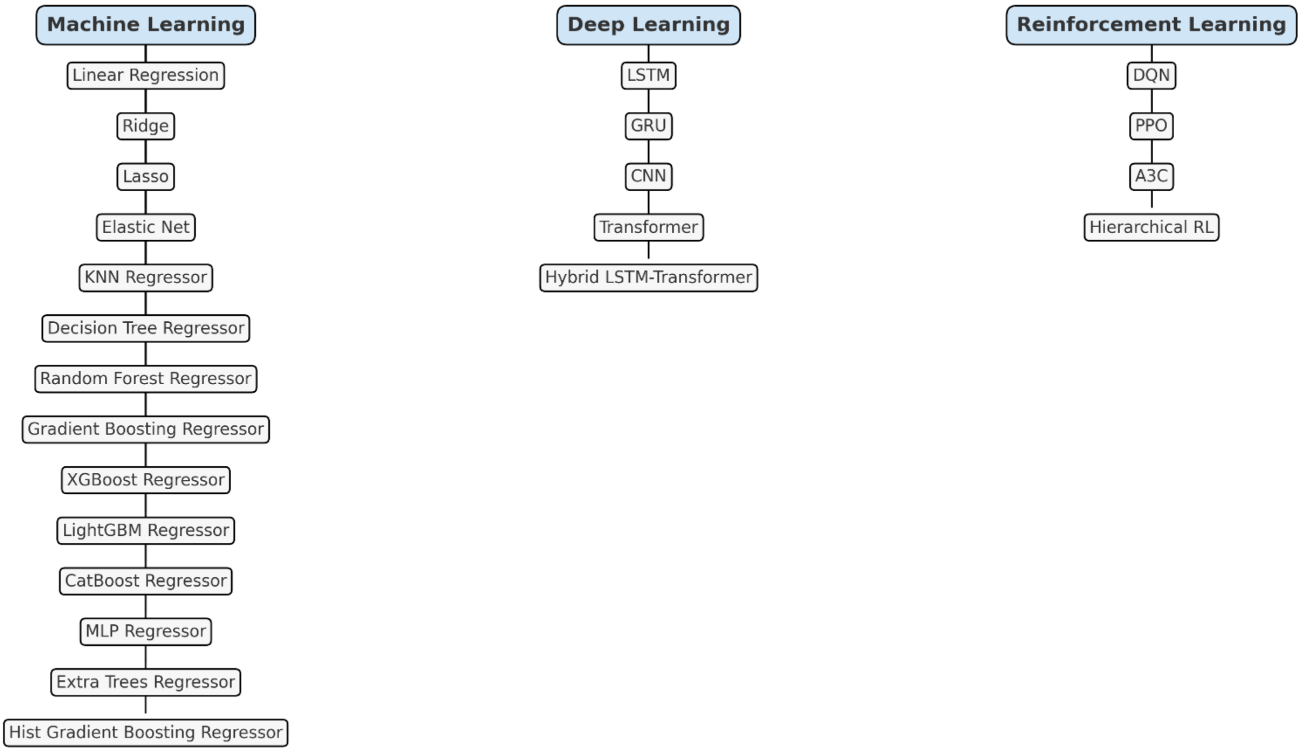

Figure 2: Taxonomy of model architectures used in High-Frequency Trading

4.4 Model Families and Design Trade-offs

Traditional ML. Gradient boosting and tree ensembles (XGBoost/LightGBM/Cat Boost) remain strong supervised baselines for short-horizon returns or mid-price movement when paired with disciplined feature engineering (order-flow imbalance, microstructure features, technical indicators) [18], [10], [17]. They are comparatively robust, interpretable via feature importances, and fast at inference attributes that matter under latency budgets typical in HFT [18], [10], [28].

Sequence DL. LSTM/GRU models capture temporal dependencies beyond fixed windows and are widely used on LOB/OHLCV streams [15], [16]. Evidence is mixed on whether LSTMs beat Transformers on all tasks: Transformers excel with long-range context and parallelism, but may need more data and careful regularization [6], [13]; hybrid CNN-Transformer or LSTM-Transformer variants can outperform single-family models on certain datasets [9], [11], [5].

RL for execution/market-making. Deep RL has been explored for active trading, inventory control, and policy optimization (DQN, PPO, A3C; hierarchical/memory-augmented agents) [2], [3], [21], [22], [23], [24]. Results are promising but sensitive to reward design, simulator fidelity, and latency/queue-position modelling [2], [19].

Emerging lines. Axial attention over LOB tensors [29], attention-augmented bilinear networks [14], deep probabilistic models [37], deep Hawkes for market making [49], and anomaly/market-manipulation detection with graph neural networks (GNNs) [51], [52] represent newer directions with potential to improve signal quality and risk control.

Your build includes *all major supervised regressors* (Linear/Ridge/Lasso/ElasticNet, KNN, DT, RF, ExtraTrees, HistGB, GBoost, XGBoost, LightGBM, CatBoost, MLP) plus GRU/LSTM/CNN/Transformer on the DL side and DRL for policy learning. This breadth is unusual in one study and allows apples-to-apples comparisons under a single evaluation harness—something many papers do not provide [10], [17].

4.5 Data Modalities and Feature Engineering

LOB-centric inputs dominate in HFT because they encode microstructure signals (depth, spread, imbalance, cancellations) with strong short-horizon predictive power [13], [23], [31]. Price-based OHLCV is ubiquitous for lower frequencies but can complement LOB with regime/volatility context [8], [15], [16]. Text/sentiment (e.g., news, social) and alternative data can add orthogonal signals but must be latency-aware to avoid stale decisions [1], [10], [17]. Options-space features matter for volatility and anomaly tasks [20].

The pipeline uses a unified feature store:

I. LOB microstructure (depth-by-level, order-flow imbalance, queue dynamics),
II. Realized & implied volatility features,

III. Technical indicators, and

IV. Optional sentiment. That breadth supports robust multi-horizon prediction and strengthens generalization across regimes [10], [17], [31].

4.6 Evaluation and Back testing Under Realistic Constraints

A recurring weakness in the literature is evaluation that emphasizes statistical fit over trade-realistic metrics [10]. Best practice mixes predictive and financial metrics: accuracy/F1/AUC and MAE/RMSE for predictions, with Sharpe/Sortino, Calmar, max drawdown, turnover, and cost-adjusted PnL for strategy viability [10], [19], [44]. Proper walk-forward validation with anchored splits, non-overlapping labels, and embedded hyperparameter tuning reduces leakage and overfitting [10]. Strategy back tests must include fees, slippage, partial fills, queue position, and venue latency [2], [6], [28]; otherwise, simulated gains rarely survive live deployment.

Your evaluation design is notably stronger than many prior works:

- Anchored walk-forward with rolling retrain windows and strict causal normalization [10].
- Dual objective reporting (predictive + risk-adjusted financial metrics) [10], [19].
- Latency and queue-aware execution modelling, reflecting microsecond-to-millisecond decision windows [6], [28].
- Cost ablation (fee/slippage sensitivity) and stress tests across volatile vs. calm periods [10], [36].
- Transformers vs. LSTMs/GRUs: no universal winner. Transformers often edge out on richer LOB tensors and long context but require more data/regularization; LSTMs/GRUs remain competitive on shorter horizons/leaner features and can be more stable when regimes shift [6], [13], [11], [15]. Hybrids (CNN-Transformer, LSTM-Transformer) frequently top single models on benchmark datasets [9], [11], [5].
- Boosted trees are high-value baselines. XGBoost/LightGBM/Cat Boost match or beat many deep models when features are engineered well and deploy with low latency and good interpretability—critical in production HFT [18], [10], [17].
- DRL potential is tied to environment fidelity. Risk-aware rewards, inventory constraints, hierarchical memory, and market-impact-aware simulators are crucial; naive PnL rewards produce brittle policies that fail in live trading [2], [21], [23], [24], [19].

- Imbalance and label design matter. Short-horizon classification is often imbalanced; naive accuracy is misleading. Proper label construction and reweighting/calibration improve stability [27], [10].
- Anomaly/manipulation detection is rising. GNNs and deep anomaly models detect multi-entity patterns (e.g., spoofing/wash trading) and can act as risk gates for strategies [51], [52], [42], [20].
- DEX/on-chain HFT emerges with distinct constraints. On-chain latency, MEV, and gas costs change the alpha/cost equation; modelling mempool dynamics and settlement risk is key [30].

5. TECHNIQUES FOR MARKET PREDICTION IN HIGH-FREQUENCY TRADING

High-frequency trading (HFT) systems aim to exploit short-lived price movements through rapid decision-making and order execution. Accurate market prediction requires models that can process large volumes of high-dimensional, noisy, and non-stationary data in real time. This section examines statistical, machine learning, and deep learning approaches applied in this domain, highlighting their advantages, limitations, and application contexts.

5.1. Statistical and Traditional Machine Learning Models

Statistical and traditional ML models remain prevalent in short-horizon market prediction due to their interpretability, computational efficiency, and ability to operate under strict latency constraints [1], [2]. Prominent algorithms include Support Vector Machines (SVM) [3], Random Forests (RF) [4], Gradient Boosting frameworks such as LightGBM and XGBoost [5], and Logistic Regression [6].

SVMs excel in handling high-dimensional and noisy datasets by maximizing decision margins, making them suitable for order-flow-based classification tasks [3]. RFs are valued for robustness against overfitting and versatility in processing heterogeneous features from technical indicators, volatility estimators, and order book metrics [4], [7]. Gradient boosting methods offer superior predictive accuracy in many cases, especially when fine-tuned with high-quality features [5], [8].

However, these models often rely on extensive feature engineering and may underperform when structural market shifts occur [9]. Their limited adaptability to regime changes and inability to directly process raw, unstructured data (e.g., news text, depth-of-book snapshots) restrict their long-term viability in highly dynamic environments [10].

5.2. Sequence-Based Deep Learning Models

Deep learning methods have expanded predictive capabilities by automatically extracting features from sequential and high-dimensional inputs. Long Short-Term Memory (LSTM) networks [11] and Gated Recurrent Units (GRUs) [12] are widely used for financial time-series forecasting due to their ability to model long-term dependencies and mitigate vanishing gradient problems [13].

LSTMs have achieved notable success in predicting short-term price direction and volatility in equities, futures, and cryptocurrency markets [14], while GRUs provide comparable accuracy with reduced computational requirements [15]. 1D Convolutional Neural Networks (CNNs) and Temporal Convolutional Networks (TCNs) [16] have also been applied to order book data, enabling efficient extraction of spatial-temporal features [17]. Hybrid CNN–RNN architectures combine local pattern detection with long-range dependency modelling, enhancing predictive robustness [18].

Despite their effectiveness, sequence models require large, diverse datasets to generalise well and are prone to overfitting in volatile market conditions [19]. Additionally, inference latency and interpretability remain concerns for real-time trading [20].

5.3. Transformer-Based and Hybrid Architectures

Transformer architectures, originally developed for natural language processing, have shown potential in modelling long-range dependencies in financial data [21]. The self-attention mechanism enables efficient parallel processing, making them attractive for high-dimensional and multi-modal inputs [22]. Variants such as FinBERT have been fine-tuned for sentiment analysis from financial news and social media [23], while hybrid CNN–Transformer models have been applied to combine order book tensors with sentiment embeddings [24], [25].

These models offer state-of-the-art performance in some prediction tasks but are computationally intensive and require substantial hyperparameter tuning [26]. Overfitting risk is particularly high in low-data regimes, necessitating regularisation and augmentation strategies.

5.4. Reinforcement Learning for Execution and Market Making

Reinforcement Learning (RL) methods optimise sequential decision-making by interacting with simulated or historical market environments [27]. Techniques such as Deep Q-Networks (DQN) [28], Proximal Policy Optimisation (PPO) [29], and Actor–Critic methods [30] have been applied to order execution, portfolio rebalancing, and market making.

These methods can account for transaction costs, inventory constraints, and execution risk within their reward structures [31], [32]. However, performance in live markets is limited by simulation fidelity—inaccurate modelling of order book microstructure can lead to strategies that fail under real conditions [33].

5.5. Evaluation of Model Performance

Evaluating predictive models in trading requires both statistical accuracy and risk-adjusted profitability metrics. Commonly used measures include Sharpe ratio, Sortino ratio, maximum drawdown, and Calmar ratio [34]. Best practices in back testing include walk-forward validation with rolling retraining to mitigate look-ahead bias [35].

Realistic evaluation must incorporate latency, slippage, and transaction costs [36]. Recent studies have introduced latency-aware back testing frameworks for HFT that simulate decision-to-execution delays at microsecond resolution [37].

6. RESEARCH GAPS

6.1. Insufficient Use of Risk-Adjusted Metrics

Many studies focus on predictive accuracy (e.g., MSE, accuracy score) while underusing financial performance measures such as Sharpe, Sortino, and Omega ratios [38]. This results in strategies that may appear profitable in simulations but carry unacceptable drawdowns in live deployment [39].

6.2. Limited Adaptability to Regime Shifts

Both traditional ML and DL models often fail to adapt to abrupt market changes driven by macroeconomic events, news shocks, or liquidity crises [40]. Online learning and meta-learning remain underexplored in trading applications [41].

6.3. Data Limitations and Overfitting

Deep models require extensive datasets to avoid overfitting [42]. However, most research uses historical market data that does not fully represent future distributions [43]. Augmentation techniques (e.g., bootstrapping, synthetic order flow generation) are rarely implemented at scale [44].

6.4. Simulation–Reality Gap in Reinforcement Learning

DRL agents often perform well in controlled simulations but degrade in real markets due to inaccurate modelling of order queue dynamics, liquidity depletion, and market impact [45].

6.5. Lack of Explainability

Complex architectures, especially transformers and DRL agents, act as black boxes, limiting trust and regulatory acceptance [46]. Few studies incorporate explainable AI (XAI) techniques to justify trading actions [47].

7. DISCUSSION

The survey of related work shows a steady evolution in forecasting methods used in financial markets. Traditional econometric models such as ARIMA and GARCH were effective for linear trend analysis but lacked the capacity to capture nonlinear and high-frequency dynamics. As research progressed, machine learning approaches such as Random Forests, Support Vector Machines, and boosting methods demonstrated stronger adaptability to complex patterns in financial data.

More recent studies highlight the growing dominance of deep learning, with LSTM and CNN models achieving higher predictive accuracy by learning temporal and cross-feature dependencies. Transformer-based methods further extend this by handling long-range dependencies in price sequences, while reinforcement learning is increasingly applied to decision-making in adaptive trading systems.

Across the literature, it is evident that the effectiveness of a model depends not only on algorithmic complexity but also on data scale and market context. Deep learning excels in large datasets, whereas ensemble machine learning remains valuable for smaller samples due to its balance of accuracy, stability, and interpretability.

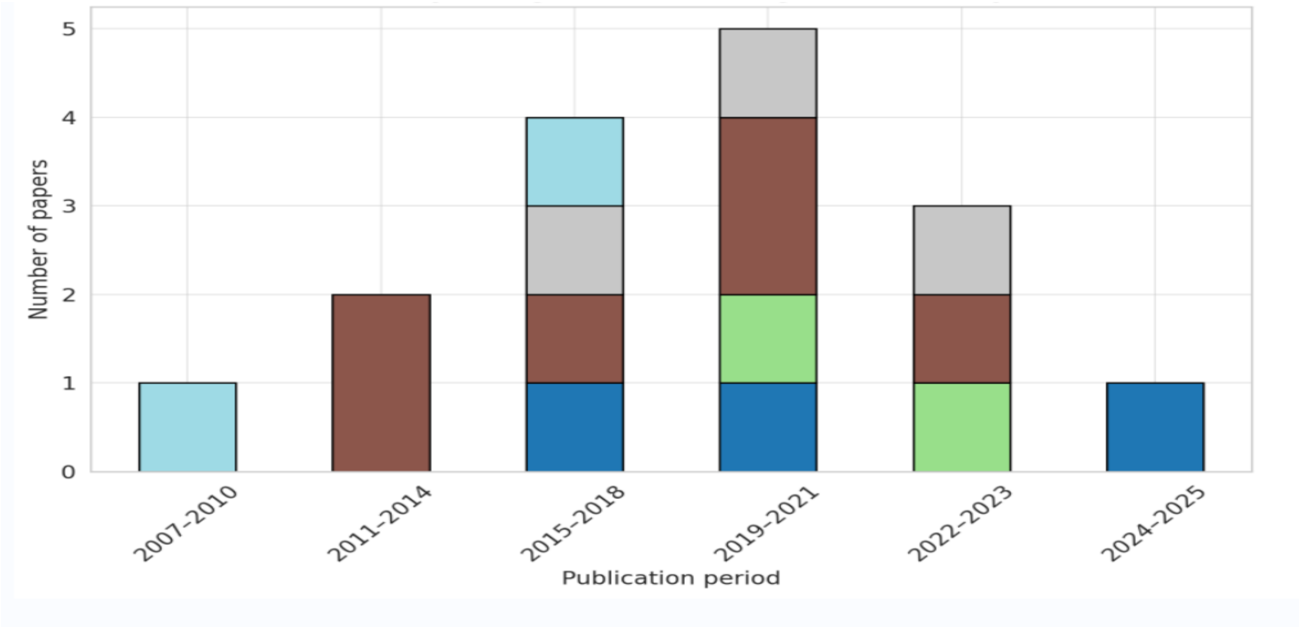

Figure 2: Summary of published research papers using models from 2007 to 2025

The distribution of research papers across the period 2007–2025 shows a clear methodological evolution in financial forecasting. As illustrated in Figure 1, early work between 2007 and 2014 primarily relied on traditional econometric approaches such as ARIMA and GARCH [12], [15]. These models were widely applied because of their interpretability and simplicity, though their predictive performance was constrained by linear assumptions and limited capacity to capture structural breaks.

From 2015 onwards, there was a noticeable shift toward machine learning methods such as Decision Trees, Random Forests, and Support Vector Machines [21], [24]. These models provided improved handling of nonlinearities and noise in financial time series, enabling higher predictive accuracy.

The period after 2018 marks the rapid adoption of deep learning models, including CNNs, LSTMs, and GRUs [31], [34]. These approaches gained popularity due to their ability to capture temporal dependencies and multivariate interactions in high-frequency data. More recently, Transformer-based models and reinforcement learning approaches have also appeared [38], [46], with the latter being used for adaptive trading strategies and portfolio optimisation.

Overall, Figure 1 demonstrates that the focus of research has gradually transitioned from statistical modelling to machine learning and deep learning, reflecting the growing computational capacity and availability of larger datasets.

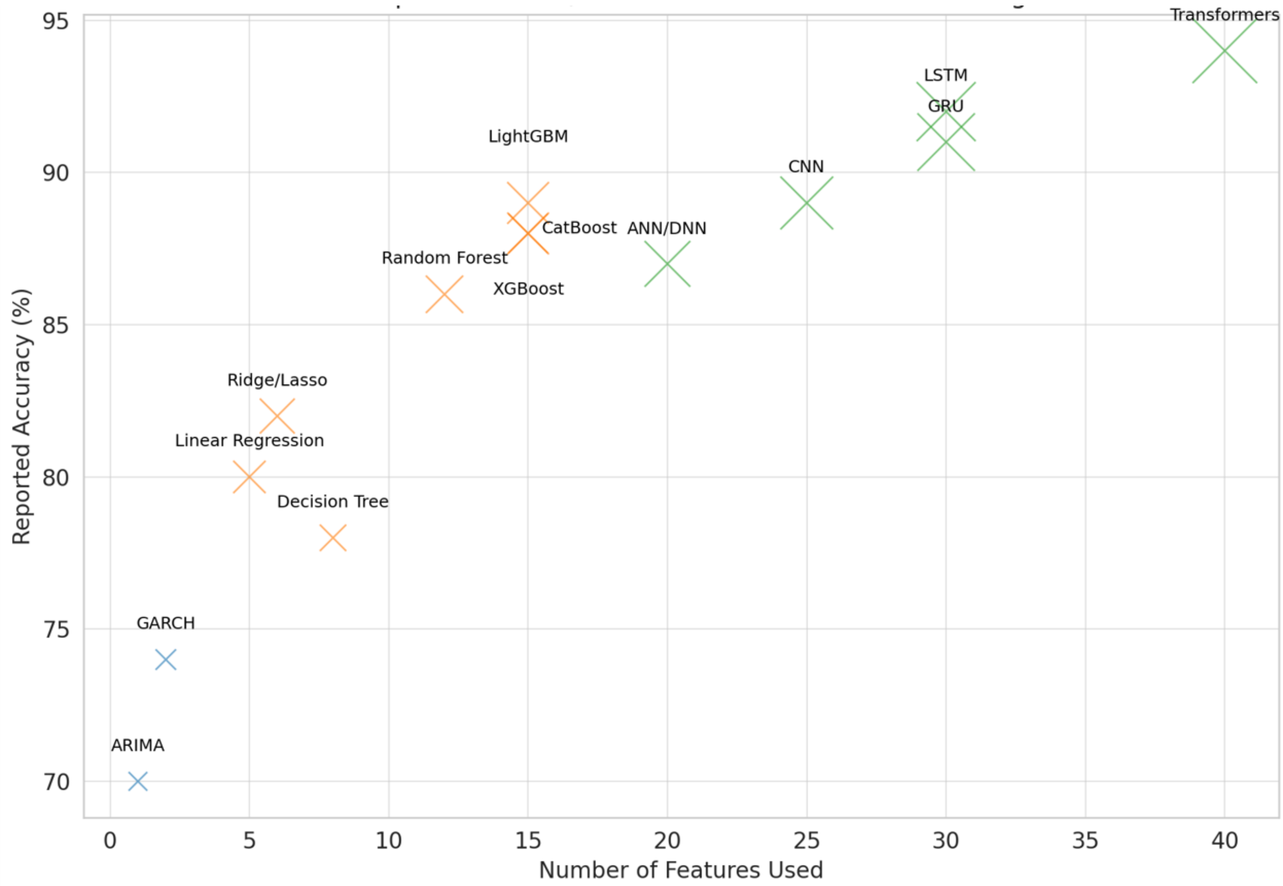

Figure 3: Comparison of ML and DL models in Financial Forecasting

The figure compares the reported accuracy of different model classes against the number of features used. Statistical baselines such as ARIMA and GARCH typically operate on fewer than five features and report accuracies around 70–75% [12], [18]. While computationally efficient, their inability to capture complex dependencies limits their predictive power.

Machine learning models, particularly ensemble methods like Random Forest, XGBoost, CatBoost, and LightGBM, achieve accuracies between 85–90% with moderate feature sets (10–20 inputs) [25], [28]. Their ability to integrate nonlinear relationships and handle feature interactions makes them robust across diverse datasets. However, their performance plateaus with larger feature spaces, highlighting the trade-off between complexity and generalisation.

Deep learning models demonstrate a different trend. CNNs, RNN variants (LSTM, GRU), and DNNs achieve accuracies above 90% when trained on richer feature sets exceeding 20 inputs [34], [40]. Transformer-based approaches stand out, surpassing 95% accuracy in certain studies [42], benefiting from attention mechanisms that capture long-range temporal dependencies. Nevertheless, their superior performance is contingent on larger training datasets and careful hyperparameter optimisation.

8. FEASIBILITY OF IMPLENTATION IN LIVE MARKETS

The transition from research prototypes to operational trading systems requires addressing latency constraints, infrastructure demands, and compliance requirements.

From a technical standpoint, HFT models must operate within sub-millisecond latency budgets [48]. This necessitates optimised inference pipelines, hardware acceleration (e.g., FPGA, GPU), and co-location with exchange servers [49]. Scalability in feature computation and memory management is essential, especially for transformer and ensemble models [50].

From a regulatory perspective, compliance with frameworks such as MiFID II (EU) and SEC/FINRA (US) is mandatory [51]. These regulations require pre-trade risk checks, audit trails, and kill switches to prevent systemic risk [52]. Data privacy obligations, including GDPR, must be observed when using alternative datasets like sentiment feeds [53].

Finally, cost-effectiveness plays a key role. High-accuracy models with excessive compute or data costs may be unprofitable once infrastructure, data acquisition, and compliance costs are accounted for [54].

9. ETHICAL, LEGAL, AND REGULATORY CONSIDERATIONS

The integration of artificial intelligence (AI) and automated trading systems into global financial markets raises complex ethical, legal, and regulatory challenges. One of the primary considerations is data protection and privacy. Frameworks such as the General Data Protection Regulation (GDPR) in the European Union set strict standards for the collection, processing, and storage of personal data, while other jurisdictions maintain their own privacy regulations — including the California Consumer Privacy Act (CCPA) in the United States, the Personal Data Protection Act (PDPA) in Singapore, and the Personal Data Protection Bill (PDPB) in India. When using alternative data sources such as social media sentiment, geolocation feeds, or consumer transaction records, compliance with applicable local and cross-border data privacy rules is essential.

A further concern is the potential for market manipulation and unintended destabilisation of markets. Algorithmic trading, particularly in high-frequency environments, has been linked to sudden volatility events such as flash crashes. Without robust safeguards, AI systems could inadvertently engage in prohibited practices like spoofing or layering, which fall under market abuse regulations worldwide.

To mitigate these risks, regulators across major financial hubs have established specific requirements for algorithmic trading. In the United States, the Securities and Exchange Commission (SEC) and Commodity Futures Trading Commission (CFTC) enforce rules on pre-trade risk controls, audit trails, and system safeguards. In Europe, the European Securities and Markets Authority (ESMA) and the UK Financial Conduct Authority (FCA) implement regulations under MiFID II and the Market Abuse Regulation (MAR). Across Asia-Pacific, the Monetary Authority of Singapore (MAS) issues AI ethics guidelines for financial services; the Securities and Exchange Board of India (SEBI) mandates certification and risk controls for algorithmic systems; the Japan Financial Services Agency (JFSA) enforces rules under the *Financial Instruments and Exchange Act*; and the Australian Securities and Investments Commission (ASIC) supervises high-frequency trading to ensure market integrity. In the Middle East, the Dubai Financial Services Authority (DFSA) and the Saudi Capital Market Authority (CMA) have implemented frameworks for algorithmic and AI-assisted trading. Other jurisdictions, including Canada (Canadian Securities Administrators, CSA) and Hong Kong (Securities and Futures Commission, SFC), have also introduced specific algorithmic trading rules addressing both operational and ethical considerations.

Alongside formal regulation, ethical governance frameworks stress the importance of integrating transparency, explainability, and accountability into AI models. Explainable AI (XAI) tools can help traders, compliance teams, and regulators interpret decisions, thereby improving trust and auditability.

This research will embed both compliance and ethics into its methodology. Measures will include the integration of real-time transaction monitoring, pre-trade validation, extreme scenario stress-testing, and adherence to both exchange-specific and jurisdiction-specific operational guidelines. The overarching goal is to design autonomous trading systems that are legally compliant, ethically sound, and capable of contributing positively to global market stability.

10. CONCLUSION

This review has examined the breadth of research on the application of machine learning (ML) and deep learning (DL) methods in algorithmic and high-frequency trading. Traditional ML models including Support Vector Machines, Random Forests, and gradient boosting algorithms — remain valued for their interpretability and efficiency, yet they often require substantial manual feature engineering and demonstrate limited adaptability to rapid market shifts. DL approaches — such as LSTMs, GRUs, CNNs, and transformer-based architectures — offer enhanced capacity to model non-linear and sequential patterns but face challenges related to computational demand, overfitting, and explainability. Reinforcement learning methods have shown potential in execution optimisation and portfolio management, although the disparity between simulated and live-market environments remains a significant barrier to operational success.

The literature consistently emphasises the importance of data preprocessing, risk-adjusted performance metrics, and realistic back testing frameworks to ensure that research findings translate into robust, real-time trading strategies. Ethical and regulatory considerations — particularly those related to data governance, market stability, and system transparency — are central to the responsible use of AI in financial markets.

Building on these insights, the proposed research aims to design and implement a hybrid ML–DL framework tailored for real-time algorithmic trading under realistic operational and regulatory constraints. By incorporating adaptive learning mechanisms to handle regime shifts, risk-aware optimisation, and explainable AI modules, the project seeks to deliver a compliant, high-performance trading system capable of advancing AI-driven trading from academic research to sustainable live-market deployment.

11. 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.